Some topics in analysis related to topological groups and Lie algebras

Stephen Semmes
Rice University
Preface
Contents

I Basic notions and objects 1

1 Absolute values and norms 2
  1.1 Metrics and ultrametrics . . . . . . . . . . . . . . . . . . . . . . . 2
  1.2 Absolute value functions . . . . . . . . . . . . . . . . . . . . . . 3
  1.3 Equivalent absolute value functions . . . . . . . . . . . . . . . . 5
  1.4 Completions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6
  1.5 Discreteness . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7
  1.6 $p$-Adic integers . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
  1.7 Residue fields . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10
  1.8 Norms and ultranorms . . . . . . . . . . . . . . . . . . . . . . . . 11
  1.9 Bounded linear mappings . . . . . . . . . . . . . . . . . . . . . . 13
  1.10 Some norms on $k^n$ . . . . . . . . . . . . . . . . . . . . . . . . . 14
  1.11 Inner products . . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
  1.12 Infinite series . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
  1.13 Bounded bilinear mappings . . . . . . . . . . . . . . . . . . . . . 20
  1.14 Minkowski functionals . . . . . . . . . . . . . . . . . . . . . . . . 22
  1.15 Balanced subgroups . . . . . . . . . . . . . . . . . . . . . . . . . 24

2 Some basic notions related to Lie algebras 26
  2.1 Modules and homomorphisms . . . . . . . . . . . . . . . . . . . . . 26
  2.2 Algebras . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27
  2.3 Lie algebras . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29
  2.4 The adjoint representation . . . . . . . . . . . . . . . . . . . . . . 31
  2.5 Derivations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32
  2.6 Involution . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34
  2.7 More on multiplication operators . . . . . . . . . . . . . . . . . . . 36
  2.8 Matrices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37
  2.9 Traces of matrices . . . . . . . . . . . . . . . . . . . . . . . . . . . 38
  2.10 Vector spaces and linear mappings . . . . . . . . . . . . . . . . . . 40
  2.11 Ideals and quotients . . . . . . . . . . . . . . . . . . . . . . . . . . 41
  2.12 Bilinear forms . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 43
  2.13 Dual spaces and mappings . . . . . . . . . . . . . . . . . . . . . . 44
  2.14 Nondegenerate bilinear forms . . . . . . . . . . . . . . . . . . . . 45
  2.15 Sesquilinear forms . . . . . . . . . . . . . . . . . . . . . . . . . . 47
CONTENTS

3 Submultiplicativity and invertibility 50
  3.1 Invertibility .......................... 50
  3.2 Submultiplicative seminorms ............ 52
  3.3 Some matrix seminorms ................. 53
  3.4 Continuity of inverses .................. 55
  3.5 Banach algebras ....................... 57
  3.6 Invertible linear mappings ............ 59
  3.7 Isometric linear mappings .............. 60
  3.8 Hilbert space isometries .............. 62
  3.9 Preserving bilinear forms ............. 63
  3.10 Preserving nondegenerate bilinear forms 65
  3.11 Preserving sesquilinear forms ......... 66
  3.12 Bilinear forms and matrices .......... 68
  3.13 Sesquilinear forms and matrices ...... 71
  3.14 Invertibility and involutions ......... 73
  3.15 Invertibility and seminorms .......... 75

4 Formal power series 77
  4.1 Direct sums and products .............. 77
  4.2 Bilinear mappings and Cauchy products 79
  4.3 Formal power series and modules ....... 81
  4.4 Sums and extensions .................. 82
  4.5 Extending bilinear mappings ........... 84
  4.6 Formal power series and algebras ...... 86
  4.7 Invertibility in $A[[T]]$ ............ 87
  4.8 Homomorphisms over $k[T]$ ............ 88
  4.9 Formal power series and homomorphisms 90
  4.10 Extensions and compositions .......... 92
  4.11 Two-step extensions ................. 94
  4.12 Extending algebra homomorphisms ...... 95
  4.13 Some remarks about commutators ...... 97
  4.14 Formal power series and involutions 98

5 Some related notions 100
  5.1 Adjoining nilpotent elements .......... 100
  5.2 Adjoining nilpotent elements to algebras 101
  5.3 Invertibility in $A[\epsilon_n]$ .......... 103
  5.4 Adjoining two nilpotent elements ...... 104
  5.5 Invertibility in $A[\epsilon_1, \eta_1]$ ...... 106
  5.6 Differentiation ...................... 107
  5.7 Polynomial functions ................. 109
  5.8 Several commuting indeterminates ...... 110
  5.9 Polynomial functions in several variables 113
  5.10 Partial derivatives .................. 114
  5.11 Formal differential operators .......... 117
  5.12 First-order differential operators ...... 118
5.13 Homogeneous formal polynomials ............... 120
5.14 Homogeneous differential operators .......... 121

6 Bilinear actions and representations .......... 123
   6.1 Bilinear actions .................................. 123
   6.2 Subactions and homomorphisms ............... 124
   6.3 Quotient actions ................................ 125
   6.4 Representations of associative algebras .... 127
   6.5 Representations of Lie algebras ............... 128
   6.6 Subrepresentations ............................... 130
   6.7 Homomorphisms between representations ..... 131
   6.8 Quotient representations ......................... 132
   6.9 Sums of representations .......................... 133
   6.10 Compatible bilinear mappings ................. 134
   6.11 Representations and formal power series ..... 136
   6.12 Opposite algebras ................................ 137
   6.13 Matrices and associative algebras .......... 139
   6.14 Irreducibility .................................... 141

7 Representations and multilinear mappings .... 142
   7.1 Some remarks about subalgebras ............... 142
   7.2 Representations on linear mappings .......... 143
   7.3 Multilinear mappings .............................. 144
   7.4 Boundedness and multilinearity ............... 146
   7.5 Representations on multilinear mappings ..... 148
   7.6 Centralizers and invariant elements .......... 149
   7.7 Invariant multilinear mappings ............... 150
   7.8 Traces of linear mappings ....................... 152
   7.9 The Killing form .................................. 154
   7.10 Invariant subspaces and traces ............... 155
   7.11 Radicals of bilinear mappings ............... 156
   7.12 Tensor products .................................. 158
   7.13 Functions on sets ................................ 159
   7.14 Some remarks about sums, products ......... 160
   7.15 Some natural isomorphisms ..................... 161

8 Formal series and ordered rings ............... 163
   8.1 Poles of finite order .............................. 163
   8.2 Sequences and series ............................. 164
   8.3 Formal series and module homomorphisms .... 166
   8.4 Extending module homomorphisms ............... 167
   8.5 Homomorphisms from $A((T))$ into $B((T))$ .... 168
   8.6 Formal series and bilinear mappings .......... 170
   8.7 Bilinear mappings on $A((T)) \times B((T))$ .... 172
   8.8 Algebras and modules over $k((T))$ .......... 173
   8.9 Absolute values on $k((T))$ ..................... 175
## CONTENTS

8.10 Formal series and algebra homomorphisms .......................... 177
8.11 Involutions and formal series ........................................ 178
8.12 Ordered rings .......................................................... 180
8.13 Some additional features .............................................. 181
8.14 Ordered fields ......................................................... 182

II Solvability, nilpotence, and semisimplicity .......................... 184

9 Solvability and nilpotence ............................................... 185
  9.1 Some basic isomorphism theorems ................................... 185
  9.2 Products of ideals ................................................... 186
  9.3 Solvable Lie algebras ................................................ 187
  9.4 The solvable radical ................................................ 188
  9.5 Nilpotent Lie algebras .............................................. 190
  9.6 Two-dimensional Lie algebras ..................................... 191
  9.7 Nilpotency conditions ............................................... 192
  9.8 Maximal Lie subalgebras ............................................ 194
  9.9 Nilpotent linear mappings .......................................... 195
  9.10 Engel’s theorem .................................................... 197
  9.11 Flags and matrices ................................................ 198
  9.12 A useful lemma .................................................... 200
  9.13 Lie’s theorem ....................................................... 202
  9.14 Structure constants ................................................ 203
  9.15 Another corollary .................................................. 204

10 Matrices and traces ...................................................... 206
  10.1 Some remarks about $gl_n(k)$ ..................................... 206
  10.2 Some basic properties of $sl_2(k)$ ............................... 208
  10.3 Scalar and diagonal matrices ..................................... 209
  10.4 Centrality in $gl_n(k), sl_n(k)$ .................................. 210
  10.5 Solvability and traces ............................................. 211
  10.6 Diagonalizable linear mappings .................................. 212
  10.7 Nilpotent vectors ................................................... 214
  10.8 Jordan–Chevalley decompositions ................................ 215
  10.9 Some related situations ........................................... 217
  10.10 Replicas .............................................................. 218
  10.11 Two lemmas about traces ........................................ 220
  10.12 Cartan’s criterion ................................................. 222
  10.13 Comparing radicals ................................................. 223
  10.14 Complementary ideals ............................................. 224
  10.15 Simple Lie algebras ............................................... 225
### 11 Some examples and related properties 227
11.1 Simplicity and solvability ................................. 227
11.2 Traces on $M_n(k)$ ........................................... 228
11.3 A nice criterion ............................................. 230
11.4 Ideals and structure constants .............................. 231
11.5 Nondegeneracy and bilinear forms ......................... 232
11.6 Bilinear forms and adjoints ................................ 233
11.7 Bilinear forms and symmetry conditions .................... 235
11.8 Traces and involutions ..................................... 236
11.9 Inner products over ordered fields ......................... 237
11.10 Self-adjoint Lie subalgebras ............................... 240
11.11 Comparing involutions .................................... 241
11.12 Some projections .......................................... 242
11.13 Antisymmetric bilinear forms .............................. 244
11.14 Symmetric bilinear forms ................................ 246

### 12 Some complex versions 248
12.1 Complexifying ordered fields ............................... 248
12.2 Sesquilinearity over $k[\sqrt{-a}]$ ......................... 249
12.3 Inner products over $k[\sqrt{-a}]$ ......................... 251
12.4 Self-adjointness in $gl(V)$ ................................ 253
12.5 Projections and sesquilinear forms ......................... 255
12.6 Nondegenerate Hermitian forms ............................ 256
12.7 Real parts .................................................. 258
12.8 Bilinear forms over $k[\sqrt{-a}]$ ......................... 260
12.9 Conjugate-linearity and adjoints ......................... 261
12.10 Antisymmetric forms over $k[\sqrt{-a}]$ .................. 262
12.11 Symmetric forms over $k[\sqrt{-a}]$ ...................... 264
12.12 Nonnegative self-adjoint operators ....................... 265
12.13 Polar decompositions ..................................... 266
12.14 Bilinear forms and inner products ....................... 268
12.15 Complex bilinear forms .................................. 269

### 13 Semisimplicity 272
13.1 Complete reducibility ..................................... 272
13.2 Casimir elements .......................................... 273
13.3 Another perspective ...................................... 275
13.4 A more particular situation ............................... 277
13.5 A splitting theorem ...................................... 277
13.6 Weyl's theorem .......................................... 279
13.7 Symmetric forms and tensors ............................. 280
13.8 Reductive Lie algebras .................................... 282
13.9 Semisimple ideals ....................................... 282
13.10 Another approach ....................................... 283
13.11 Inner derivations ....................................... 285
14 Nilpotence and diagonalizability 286
14.1 A duality argument 286
14.2 A criterion for nilpotence 287
14.3 ad-Diagonalizability 288
14.4 Semisimple subalgebras of \( gl(V) \) 290
14.5 Tensors 291
14.6 Induced linear mappings 292
14.7 Invariant tensors 294
14.8 Diagonalizability and quotients 294
14.9 Diagonalizability and representations 296
14.10 Exponentiating nilpotents 297
14.11 Nilpotent derivations 299
14.12 Exponentials and homomorphisms 301
14.13 Representations and structure constants 302
14.14 Lie’s theorem and nilpotence 303

15 Complexifications and \( sl_2(k) \) modules 305
15.1 Basic properties of \( sl_2(k) \) modules 305
15.2 Maximal or primitive vectors 306
15.3 Related submodules 307
15.4 Constructing modules \( W(m) \) 309
15.5 More on \( W(m) \) 310
15.6 Another construction 312
15.7 Some exponentials 313
15.8 Exponentials and \( W(m) \) 315
15.9 Additional properties of \( sl_2(k) \) modules 316
15.10 Complexifying real vector spaces 317
15.11 Complexifying \( sl_2(k) \) modules 318
15.12 Complexifying algebras over \( R \) 320
15.13 Conjugate-linear involutions 321

16 Module homomorphisms 324
16.1 Bilinear actions and homomorphisms 324
16.2 Homomorphisms and semisimplicity 325
16.3 Linear mappings and dimensions 327
16.4 Homomorphisms and dimensions 327
16.5 Semisimplicity and dimensions 328
16.6 Absolutely simple Lie algebras 329
16.7 Algebras over subrings 331
16.8 Lie algebras over subfields 332

III Toral subalgebras and roots 334
17 Subalgebras and diagonalizability 335
17.1 Toral subalgebras 335
17.2 Simultaneous diagonalizability .................................. 336
17.3 Related bilinear forms .............................................. 338
17.4 Maximal toral subalgebras .......................................... 339
17.5 Self-centralizability and diagonalizability ....................... 341
17.6 Characteristic 0, \( Z(A) = \{0\} \) ................................. 343
17.7 The dimension of \( A_{\alpha} \) .......................................... 344
17.8 Other elements of \( \Phi_B \) ............................................. 347
17.9 Adding elements of \( \Phi_B \) ............................................. 349
17.10 The dual of \( B \) .................................................... 351
17.11 Using the Killing form .............................................. 353
17.12 Vector spaces over \( \mathbb{Q} \) .......................................... 354
17.13 Vector spaces over \( \mathbb{R} \) .......................................... 355
17.14 Normalizers and centralizers ...................................... 357
17.15 Semisimplicity and diagonalizability ............................ 358

18 Cartan subalgebras .................................................... 360
18.1 Subspaces related to \( ad_x \) ........................................... 360
18.2 Some remarks about characteristic polynomials .............. 361
18.3 Characteristic polynomials and \( ad_x \) ............................. 362
18.4 Engel subalgebras .................................................. 363
18.5 Self-normalizing subalgebras ...................................... 365
18.6 Orthogonality conditions ........................................... 366
18.7 Nilpotence and normalizing elements ............................ 367
18.8 Minimal Engel subalgebras ........................................ 368
18.9 Nilpotent vectors and sums ....................................... 369
18.10 Cartan subalgebras and semisimplicity ......................... 371
18.11 Subalgebras and homomorphisms ................................ 373
18.12 Unions of hyperplanes ........................................... 374
18.13 Centralizers and diagonalizability ............................... 375

19 Root systems ............................................................ 377
19.1 Symmetries and reflections ....................................... 377
19.2 Invariant inner products .......................................... 378
19.3 Defining root systems ............................................. 379
19.4 The Weyl group .................................................... 381
19.5 Isomorphic root systems .......................................... 382
19.6 Duality and symmetries .......................................... 383
19.7 Ad joints and reflections ......................................... 385
19.8 Inverse roots ........................................................ 387
19.9 Angles between roots ............................................. 389
19.10 A criterion for linear independence ............................ 391
19.11 Bases for root systems ........................................... 392
19.12 Positive roots ..................................................... 394
19.13 Bases and duality ................................................ 396
19.14 Bases and the Weyl group ....................................... 398
19.15 Weyl chambers ................................................... 401
## CONTENTS

22.12 Standard Borel subalgebras ........................................... 472
22.13 Subalgebras and derivations ........................................... 473
22.14 Reducibility and semigroups ............................................ 474
22.15 Reducibility and roots .................................................. 477

23 Roots and Lie algebras, 2 .................................................. 479
  23.1 Inverse roots and homomorphisms .................................... 479
  23.2 Bases and submodules .................................................. 480
  23.3 Some generating submodules .......................................... 481
  23.4 Some generators ....................................................... 483
  23.5 Some relations ........................................................ 484
  23.6 Isomorphisms and diagonalizability ................................ 486
  23.7 Automorphisms and sl$_2(k)$ ......................................... 487
  23.8 Automorphisms and roots ............................................. 488
  23.9 Automorphisms and bases for $B$ .................................... 489
  23.10 Sums and diagonalizability ......................................... 490

24 Strong $\text{ad}$-nilpotence and automorphisms .......................... 498
  24.1 Invariant subspaces and nilpotent vectors ......................... 498
  24.2 Finite-dimensional spaces .......................................... 499
  24.3 Strong $\text{ad}$-nilpotence ............................................ 500
  24.4 Strong $\text{ad}$-nilpotence and homomorphisms ..................... 501
  24.5 Another criterion for nilpotence ................................... 502
  24.6 Nilpotence and automorphisms ...................................... 503
  24.7 Conjugacy and solvability .......................................... 504
  24.8 The case where $A = B_2$ ........................................... 505
  24.9 Conjugacy in arbitrary Lie algebras ................................ 508
    24.9.1 The beginning of the proof of the theorem ................... 508
    24.9.2 Standard Borel subalgebras of $A$ ............................. 509
    24.9.3 The second induction hypothesis .............................. 510
  24.10 The first subcase .................................................. 510
    24.10.1 Nilpotent elements of $B_\Delta \cap B$ ....................... 511
    24.10.2 Some Borel subalgebras of $M$ and $A$ ..................... 513
  24.11 The second subcase ................................................ 514
    24.11.1 Reducing to $T \subseteq A_0$ .................................... 515
    24.11.2 The case where $T = A_0$ ..................................... 516
    24.11.3 The case where $T \neq A_0$, $B \subseteq C_A(T)$ ............ 516
    24.11.4 The case where $T \neq A_0$, $B \nsubseteq C_A(T)$ ......... 517
  24.12 The second case .................................................... 518
  24.13 Some consequences .................................................. 520
  24.14 The case of sl$_2(k)$ ............................................... 522
  24.15 Automorphisms and semisimplicity ................................. 523
IV Some related constructions 525

25 Universal enveloping algebras 526
  25.1 Polynomial algebras 526
  25.2 Tensor products of associative algebras 528
  25.3 Tensor algebras 529
  25.4 Defining universal enveloping algebras 530
  25.5 Constructing universal enveloping algebras 531
  25.6 Some properties of $U\mathcal{A}$ 532
  25.7 Commuting Lie subalgebras 533
  25.8 Symmetric algebras 535
  25.9 A filtration on $U\mathcal{A}$ 537
  25.10 The Poincaré-Birkhoff-Witt theorem 539
  25.11 The rest of the proof 541
  25.12 Some corollaries 545
  25.13 Some related filtrations 546
  25.14 The diagonal mapping 548
  25.15 More on primitive elements 550

26 Free Lie algebras 554
  26.1 Magmas and their homomorphisms 554
  26.2 Magmas and algebras 555
  26.3 Constructing free Lie algebras 555
  26.4 Degree 2 in $L_X$ 557
  26.5 Free associative algebras 559
  26.6 Injectivity and free modules 560
  26.7 Finite sets $X, k = \mathbb{Z}$ 562
  26.8 Arbitrary sets $X, k = \mathbb{Z}$ 563
  26.9 Arbitrary $X, k$ 564
  26.10 Bases from P. Hall families 565
  26.11 More on $L^r_X$ 566

27 Some more Lie algebras 567
  27.1 Generators and relations 567
  27.2 A module over $\mathcal{L}$ 568
  27.3 Some properties of $\mathcal{L}$ 572
  27.4 Some brackets in $\mathcal{L}$ 574
  27.5 Eigenvectors in $\mathcal{L}$ 575
  27.6 Some more properties of $\mathcal{L}$ 576
  27.7 Some more relations in $\mathcal{L}$ 578
  27.8 Serre’s theorem 579
  27.9 Ideals $\mathcal{I}, \mathcal{J}$ 580
  27.10 Subalgebras $\mathcal{X}, \mathcal{H}$, and $\mathcal{Y}$ of $L$ 581
  27.11 Eigenvectors in $L$ 582
  27.12 Locally nilpotent linear mappings 584
  27.13 Some properties of $\theta_\alpha$ 585
## CONTENTS

27.14 Roots and $\mathcal{H}'$ ........................................... 587  
27.15 Some dimensions of eigenspaces ............................... 588  

### 28 Some more Lie algebras, 2

28.1 Linear combinations of roots ................................... 590  
28.2 The dimension of $L$ .............................................. 591  
28.3 Automorphisms and the Weyl group ............................ 592  
28.4 Semisimplicity of $L$ ............................................ 594  
28.5 An isomorphism theorem ......................................... 595  
28.6 Some automorphisms ............................................. 596  
28.7 Roots and diagonalizability ..................................... 598  
28.8 Quotients of $L$ .................................................. 599  
28.9 Additional identities in $B$ ..................................... 600  
28.10 Eigenspaces in $L$ ............................................... 601  
28.11 Subsets of $\Delta$ ............................................... 603  

### 29 Chevalley’s normalization

29.1 Linearly independent roots ..................................... 605  
29.2 An identity for root strings .................................... 607  
29.3 Related properties of $\lambda_\alpha$’s ............................. 609  
29.4 Associated Lie algebras ......................................... 609  
29.5 A lemma about brackets ........................................ 611  
29.6 Some remarks about automorphisms ............................ 612  
29.7 Diagonal automorphisms ........................................ 613  
29.8 Basis vectors in $A$ ................................................ 614  
29.9 Some properties of the $c_{\alpha,\beta}$’s ......................... 616  
29.10 Uniqueness of Chevalley bases ................................ 617  
29.11 The complex case ............................................... 619  
29.12 The corresponding Killing forms ............................... 621  
29.13 Using other commutative rings ................................ 624  
29.14 Some mappings on $A(\mathbb{Z})$ ............................... 625  
29.15 Some related mappings on $A(k_1)$ ........................... 627  

### 30 Roots and abstract weights

30.1 Abstract weights .................................................. 630  
30.2 Some remarks about lattices .................................... 631  
30.3 Related positivity conditions ................................... 633  
30.4 Automorphisms and $\Upsilon$ .................................... 635  
30.5 Some partial orderings on $V$ ................................... 637  
30.6 Some more properties of $\Upsilon$ ............................... 638  
30.7 Saturated subsets of $\Upsilon$ ................................... 640  
30.8 More on highest weights ........................................ 641  
30.9 More on saturated sets .......................................... 643  
30.10 Another look at saturated sets ................................. 644  
30.11 A property of saturated sets ................................. 646
V SOME REPRESENTATION THEORY 647

31 REPRESENTATIONS AND SEMISIMPLICITY 648
31.1 Diagonalizability and roots ................................. 648
31.2 Using a larger field ........................................... 649
31.3 Related spaces of linear mappings .......................... 650
31.4 Larger fields and toral subalgebras .......................... 652
31.5 Weights and modules ......................................... 655
31.6 Primitive or maximal vectors .................................. 656
31.7 Representations and UA ...................................... 658
31.8 Spanning standard cyclic modules ............................ 658
31.9 Indecomposability ............................................. 661
31.10 Irreducible standard cyclic modules ......................... 663
31.11 Representations and left ideals ............................... 665
31.12 Tensor products over algebras ............................... 666
31.13 Existence of standard cyclic modules ....................... 668

32 REPRESENTATIONS AND SEMISIMPLICITY, 2 671
32.1 Weights and finite dimension .................................. 671
32.2 Weights and a larger field .................................... 673
32.3 Conditions for finite dimension ............................... 674
32.4 Connections with abstract weights ............................ 676
32.5 Some initial steps .............................................. 677
32.6 Submodules over subalgebras ................................. 680
32.7 Some related exponentials on V ............................... 681
32.8 Weights and the Weyl group .................................. 683
32.9 Sums of weights ............................................... 684
32.10 Weights and tensor products ................................. 686
32.11 Fundamental weights and modules ........................... 687
32.12 Simple transitivity ........................................... 688
32.13 Weights and saturated sets .................................. 689

BIBLIOGRAPHY 692

INDEX 694
Part I

Basic notions and objects
Chapter 1

Absolute values and norms

1.1 Metrics and ultrametrics

Let $X$ be a set. A nonnegative real-valued function $d(x, y)$ defined for $x, y \in X$ is said to be a semimetric on $X$ if it satisfies the following three conditions. First,
\begin{equation}
(1.1.1) \quad d(x, x) = 0
\end{equation}
for every $x \in X$. Second,
\begin{equation}
(1.1.2) \quad d(x, y) = d(y, x)
\end{equation}
for every $x, y \in X$. Third,
\begin{equation}
(1.1.3) \quad d(x, z) \leq d(x, y) + d(y, z)
\end{equation}
for every $x, y, z \in X$. If we also have that
\begin{equation}
(1.1.4) \quad d(x, y) > 0
\end{equation}
for every $x, y \in X$ with $x \neq y$, then $d(\cdot, \cdot)$ is said to be a metric on $X$. The discrete metric is defined on $X$ by putting $d(x, y)$ equal to 0 when $x = y$, and equal to 1 when $x \neq y$.

Similarly, a nonnegative real-valued function $d(x, y)$ defined for $x, y \in X$ is said to be a semi-ultrametric on $X$ if if it satisfies (1.1.1), (1.1.2), and
\begin{equation}
(1.1.5) \quad d(x, z) \leq \max(d(x, y), d(y, z))
\end{equation}
for every $x, y, z \in X$. Note that (1.1.5) implies (1.1.3), so that a semi-ultrametric on $X$ is a semimetric on $X$ in particular. If a semi-ultrametric $d(x, y)$ on $X$ satisfies (1.1.4), then $d(x, y)$ is said to be an ultrametric on $X$. It is easy to see that the discrete metric on $X$ is an ultrametric.

Let $d(x, y)$ be a semimetric on $X$. The open ball in $X$ centered at $x \in X$ with radius $r > 0$ with respect to $d(\cdot, \cdot)$ is defined as usual by
\begin{equation}
(1.1.6) \quad B(x, r) = B_d(x, r) = \{y \in X : d(x, y) < r\}.
\end{equation}
1.2. ABSOLUTE VALUE FUNCTIONS

Similarly, the closed ball in $X$ centered at $x \in X$ with radius $r \geq 0$ is defined by

$$B(x, r) = \overline{B}(x, r) = \{ y \in X : d(x, y) \leq r \}. \tag{1.1.7}$$

An open set $U$ of $X$ is said to be an open set with respect to $d(\cdot, \cdot)$ if for every $x \in U$ there is an $r > 0$ such that

$$B(x, r) \subseteq U. \tag{1.1.8}$$

This defines a topology on $X$, by standard arguments. One can check that open balls in $X$ are open sets, and that closed balls are closed sets. If $d(\cdot, \cdot)$ is a metric on $X$, then $X$ is Hausdorff with respect to the topology determined by $d(\cdot, \cdot)$.

Suppose that $d(\cdot, \cdot)$ is a semi-ultrametric on $X$. If $x, y \in X$ satisfy $d(x, y) < r$ for some $r > 0$, then it is easy to see that

$$B(x, r) \subseteq B(y, r). \tag{1.1.9}$$

More precisely,

$$B(x, r) = B(y, r), \tag{1.1.10}$$

because we can interchange the roles of $x$ and $y$ in (1.1.9). Similarly, if $x, y \in X$ satisfy $d(x, y) \leq r$ for some $r \geq 0$, then

$$\overline{B}(x, r) \subseteq \overline{B}(y, r), \tag{1.1.11}$$

and hence

$$\overline{B}(x, r) = \overline{B}(y, r). \tag{1.1.12}$$

This implies that closed balls in $X$ with positive radius are open sets, and one can check that open balls in $X$ are closed sets in this situation.

1.2 Absolute value functions

Let $k$ be a field. A nonnegative real-valued function $|\cdot|$ on $k$ is said to be an absolute value function on $k$ if it satisfies the following conditions. First, $|x| = 0$ if and only if $x = 0$. Second,

$$|xy| = |x||y| \tag{1.2.1}$$

for every $x, y \in k$. Third,

$$|x + y| \leq |x| + |y| \tag{1.2.2}$$

for every $x, y \in k$. The standard absolute value functions on the fields $\mathbb{R}$ of real numbers and $\mathbb{C}$ of complex numbers are absolute value functions in this sense. The trivial absolute value function on any field $k$ is defined by putting $|x|$ equal to 0 when $x = 0$, and equal to 1 when $x \neq 0$.

If $|\cdot|$ is any absolute value function on a field $k$, then $|1| = 1$, where the first 1 is the multiplicative identity element in $k$, and the second 1 is the multiplicative identity element in $\mathbb{R}$. This uses the fact that $1^2 = 1$ in $k$, so that $|1| = |1|^2$ by
CHAPTER 1. ABSOLUTE VALUES AND NORMS

(1.2.1) If \( x \in k \) satisfies \( x^n = 1 \) for some positive integer \( n \), then \( |x|^n = |1| = 1 \), and hence \( |x| = 1 \). In particular, \( |-1| = 1 \), because \((-1)^2 = 1\). It follows that

(1.2.3) \[ d(x, y) = |x - y| \]

defines a metric on \( k \), using \( | -1| = 1 \) to get that (1.2.3) is symmetric in \( x \) and \( y \).

A nonnegative real-valued function \( |\cdot| \) on a field \( k \) is said to be an ultrametric absolute value function on \( k \) if it satisfies the first two conditions in the definition of an absolute value function, and

(1.2.4) \[ |x + y| \leq \max(|x|, |y|) \]

for every \( x, y \in k \). It is easy to see that (1.2.4) implies (1.2.2), so that an ultrametric absolute value function on \( k \) is an absolute value function on \( k \). If \( |\cdot| \) is an ultrametric absolute value function on \( k \), then (1.2.3) is an ultrametric on \( k \). The trivial absolute value function on any field \( k \) is an ultrametric absolute value function. The ultrametric associated to the trivial absolute value function as in (1.2.3) is the discrete metric.

Let \( p \) be a prime number. The \( p \)-adic absolute value \( |x|_p \) of a rational number \( x \) is defined as follows. If \( x = 0 \), then we put \( |x|_p = 0 \). Otherwise, if \( x \neq 0 \), then \( x \) can be expressed as \( p^j (a/b) \) for some integers \( a, b \), and \( j \), where \( a, b \neq 0 \), and neither \( a \) nor \( b \) is an integer multiple of \( p \). In this case, we put

(1.2.5) \[ |x|_p = p^{-j}. \]

One can check that this defines an ultrametric absolute value function on the field \( \mathbb{Q} \) of rational numbers. The corresponding ultrametric

(1.2.6) \[ d_p(x, y) = |x - y|_p \]

is known as the \( p \)-adic metric on \( \mathbb{Q} \).

Let \( k \) be any field again, and let \( \mathbb{Z}_+ \) be the set of positive integers, as usual. If \( x \in k \) and \( n \in \mathbb{Z}_+ \), then let \( n \cdot x \) be the sum of \( n \) \( x \)'s in \( k \). An absolute value function \( |\cdot| \) on \( k \) is said to be archimedean on \( k \) if there are \( n \in \mathbb{Z}_+ \) such that \( |n \cdot 1| \) is arbitrarily large. Otherwise, \( |\cdot| \) is said to be non-archimedean on \( k \). If \( |\cdot| \) is an ultrametric absolute value function on \( k \), then it is easy to see that

(1.2.7) \[ |n \cdot 1| \leq 1 \]

for every \( n \in \mathbb{Z}_+ \), so that \( |\cdot| \) is non-archimedean on \( k \). Conversely, it is well known that a non-archimedean absolute value function on \( k \) is necessarily an ultrametric absolute value function on \( k \). In particular, (1.2.7) holds for every \( n \in \mathbb{Z}_+ \) in this case, which can be verified more directly. More precisely, if \( |\cdot| \) is any absolute value function on \( k \), then one can check that

(1.2.8) \[ |n^j \cdot 1| = |(n \cdot 1)^j| = |n \cdot 1|^j \]

for all positive integers \( j, n \). If \( |n \cdot 1| > 1 \) for some \( n \in \mathbb{Z}_+ \), then (1.2.8) tends to \( +\infty \) as \( j \to \infty \), so that \( |\cdot| \) is archimedean on \( k \).
1.3 Equivalent absolute value functions

If \( a \) is a positive real number with \( a \leq 1 \), then it is well known that
\[
(r + t)^a \leq r^a + t^a
\]
for all nonnegative real numbers \( a, b \). To see this, observe first that
\[
\max(r, t) \leq (r^a + t^a)^{1/a}
\]
for every \( a > 0 \). We also have that
\[
(r + t)^{1-a} r^a + t^a \leq \max(r^{1-a}, t^{1-a}) (r^a + t^a).
\]
If \( a \leq 1 \), then it follows that
\[
(r + t)^{1-a} r^a + t^a \leq \max(r^{1-a}, t^{1-a}) (r^a + t^a) = (r^a + t^a)^{1/a},
\]
using (1.3.2) in the second step. This implies (1.3.1), as desired.

Let \( d(x, y) \) be a semimetric on a set \( X \). If \( 0 < a \leq 1 \), then one can check that
\[
d(x, y)^a
\]
also defines a semimetric on \( X \). More precisely, one can verify that (1.3.5) satisfies the triangle inequality using (1.3.1) and the triangle inequality for \( d(x, y) \). If \( d(x, y) \) is a semi-ultrametric on \( X \), then (1.3.5) is a semi-ultrametric on \( X \) for every \( a > 0 \).

Suppose that \( d(x, y) \) is a semimetric on \( X \) again, and that (1.3.5) is a semimetric on \( X \) too for some \( a > 0 \). Observe that
\[
B_{d^a}(x, r^a) = B_d(x, r)
\]
for every \( x \in X \) and \( r > 0 \), and that
\[
\overline{B}_{d^a}(x, r^a) = \overline{B}_d(x, r)
\]
for every \( x \in X \) and \( r \geq 0 \). In particular, (1.3.6) implies that \( d(x, y) \) and (1.3.5) determine the same topology on \( X \).

Let \( k \) be a field, and let \( | \cdot | \) be an absolute value function on \( k \). If \( 0 < a \leq 1 \), then \( |x|^a \) also defines an absolute value function on \( k \). As before, this uses (1.3.1) to get the triangle inequality for \( |x|^a \) from the triangle inequality for \( | \cdot | \). If \( | \cdot | \) is an ultrametric absolute value function on \( k \), then \( |x|^a \) is an ultrametric absolute value function on \( k \) for every \( a > 0 \).

Suppose that \( | \cdot | \) is an absolute value function on \( k \) again, and that \( | \cdot |^a \) is an absolute value function on \( k \) as well for some \( a > 0 \). Thus the metric associated to \( | \cdot |^a \) on \( k \) is the same as the \( a \)th power of the metric associated to \( | \cdot | \) on \( k \). Hence these two metrics determine the same topology on \( k \), as in the preceding paragraph.
CHAPTER 1. ABSOLUTE VALUES AND NORMS

Let $|\cdot|_1$ and $|\cdot|_2$ be absolute value functions on $k$. If there is a positive real number $a$ such that

$$|x|_2 = |x|_1^a$$

for every $x \in k$, then $|\cdot|_1$ and $|\cdot|_2$ are said to be equivalent on $k$. In this case, the metrics associated to $|\cdot|_1$ and $|\cdot|_2$ determine the same topology on $k$, as in the previous paragraph. Conversely, if the metrics associated to $|\cdot|_1$ and $|\cdot|_2$ determine the same topology on $k$, then it is well known that $|\cdot|_1$ and $|\cdot|_2$ are equivalent on $k$, in the sense of (1.3.8).

Let $|\cdot|$ be an absolute value function on $\mathbb{Q}$. A famous theorem of Ostrowski implies that $|\cdot|$ is either equivalent to the standard absolute value function on $\mathbb{Q}$, or $|\cdot|$ is the trivial absolute value function on $\mathbb{Q}$, or $|\cdot|$ is equivalent to the $p$-adic absolute value function on $\mathbb{Q}$ for some prime number $p$.

1.4 Completions

Let $(X,d_X)$ and $(Y,d_Y)$ be metric spaces, and let $E$ be a dense subset of $X$. Suppose that $f$ is a uniformly continuous mapping from $E$ into $Y$, with respect to the restriction of $d_X$ to $E$. If $Y$ is complete with respect to $d_Y$, then it is well known that there is a unique extension of $f$ to a uniformly continuous mapping from $X$ into $Y$. More precisely, uniqueness only uses continuity of the extension.

If $X$ is not complete, then it is well known that one can pass to a completion, which is given by an isometric mapping from $X$ onto a dense subset of a complete metric space. The completion is unique up to isometric equivalence, because of the extension theorem mentioned in the preceding paragraph.

Let $X$ be a set with a semimetric $d(x,y)$, and let $E$ be a dense subset of $X$. If the restriction of $d(x,y)$ to $x,y \in E$ defines a semi-ultrametric on $E$, then one can check that $d(x,y)$ is a semi-ultrametric on $X$. In particular, the completion of an ultrametric space is an ultrametric space too.

Let $k$ be a field, and let $|\cdot|$ be an absolute value function on $k$. If $k$ is not complete with respect to the metric associated to $|\cdot|$, then one can pass to a completion. It is well known that the field operations on $k$ can be extended to the completion, in such a way that the completion is also a field. The absolute value function on $k$ can be extended to an absolute value function on the completion, which corresponds to the distance to 0 in the completion. The completion of $k$ is unique, up to isometric isomorphic equivalence.

If $|\cdot|$ is an ultrametric absolute value function on $k$, then the extension of $|\cdot|$ to the completion of $k$ is an ultrametric absolute value function as well. This is analogous to the earlier statement for ultrametric spaces, and can be obtained from that statement. Alternatively, let $k_1$ be any field with an absolute value function $|\cdot|$, and let $k_0$ be a subfield of $k_1$. It is easy to see that $|\cdot|$ is archimedean on $k_1$ if and only if the restriction of $|\cdot|$ to $k_0$ is archimedean on $k_0$.

Let $p$ be a prime number. The field $\mathbb{Q}_p$ of $p$-adic numbers is obtained by completing $\mathbb{Q}$ with respect to the $p$-adic absolute value function $|\cdot|_p$. The corresponding extension of $|\cdot|_p$ to $\mathbb{Q}_p$ is also denoted $|\cdot|_p$, and defines an
ultrametric absolute value function on $\mathbf{Q}_p$. If $x \in \mathbf{Q}_p$ and $x \neq 0$, then one can check that $|x|_p$ is an integer power of $p$.

Let $k$ be a field with an absolute value function $| \cdot |$ again. If $k$ has positive characteristic, then it is easy to see that $| \cdot |$ is non-archimedean on $k$. Suppose that $| \cdot |$ is archimedean on $k$, which implies that $k$ has characteristic 0. If $k$ is complete with respect to the metric associated to $| \cdot |$, then another famous theorem of Ostrowski implies that $k$ is isomorphic to $\mathbf{R}$ or $\mathbf{C}$, in such a way that $| \cdot |$ corresponds to an absolute value function on $\mathbf{R}$ or $\mathbf{C}$ that is equivalent to the standard absolute value function.

1.5 Discreteness

Let $k$ be a field, and let $| \cdot |$ be an absolute value function on $k$. Observe that

$$(1.5.1) \quad \{ |x| : x \in k, x \neq 0 \}$$

is a subgroup of the group $\mathbf{R}_+$ of positive real numbers with respect to multiplication. Of course, (1.5.1) is the trivial subgroup $\{1\}$ of $\mathbf{R}_+$ exactly when $| \cdot |$ is the trivial absolute value function on $k$. If 1 is not a limit point of (1.5.1) with respect to the standard topology on $\mathbf{R}$, then $| \cdot |$ is said to be discrete on $k$.

Put

$$(1.5.2) \quad \rho_1 = \sup \{|x| : x \in k, |x| < 1\},$$

so that $0 \leq \rho_1 \leq 1$. If $| \cdot |$ is the trivial absolute value function on $k$, then $\rho_1 = 0$. Conversely, if $| \cdot |$ is not the trivial absolute value function on $k$, then there is a $y \in k$ such that $y \neq 0$ and $|y| \neq 1$. This implies that there is an $x \in k$ such that $x \neq 0$ and $|x| < 1$, by taking $x = y$ when $|y| < 1$ and $x = 1/y$ when $|y| > 1$. Thus $\rho_1 > 0$ when $| \cdot |$ is nontrivial on $k$.

If $| \cdot |$ is a discrete absolute value function on $k$, then $\rho_1 < 1$. Conversely, if $\rho_1 < 1$, then $| \cdot |$ is discrete on $k$. More precisely, the definition of $\rho_1$ implies that there is no $x \in k$ such that $\rho_1 < |x| < 1$. If $y \in k$ and $|y| > 1$, then we can apply the previous statement to $x = 1/y$, to get that $1/|y| \leq \rho_1$. It follows that 1 is not a limit point of (1.5.1) in $\mathbf{R}$ when $\rho_1 < 1$, as desired.

Suppose that $| \cdot |$ is an archimedean absolute value function on $k$. This implies that $k$ has characteristic 0, as in the previous section. Hence there is a natural embedding of $\mathbf{Q}$ into $k$. This leads to an absolute value function on $\mathbf{Q}$, using $| \cdot |$ on $k$. It is easy to see that $\mathbf{Q}$ is archimedean with respect to this absolute value function, because $k$ is archimedean with respect to $| \cdot |$. Using this and Ostrowski’s classification of absolute value functions on $\mathbf{Q}$ mentioned in Section 1.3, we get that this absolute value function on $\mathbf{Q}$ is equivalent to the standard absolute value function. In particular, it follows that this absolute value function on $\mathbf{Q}$ is not discrete. This means that $| \cdot |$ is not discrete on $k$. If $| \cdot |$ is a discrete absolute value function on $k$, then $| \cdot |$ is non-archimedean on $k$, and hence $| \cdot |$ is an ultrametric absolute value function on $k$. 
CHAPTER 1. ABSOLUTE VALUES AND NORMS

Suppose that $|\cdot|$ is a nontrivial discrete absolute value function on $k$, so that $0 < \rho_1 < 1$. If $y, z \in k$ and $|y| < |z|$, then

$$|y| \leq \rho_1 |z|, \quad (1.5.3)$$

because $|y/z| = |y|/|z| < 1$, and hence $|y/z| \leq \rho_1$, by the definition (1.5.2) of $\rho_1$. One can check that the supremum is attained in (1.5.2), since otherwise there would be distinct elements of (1.5.1) close to $1$, whose quotient would be close to 1 but not equal to 1. Thus $\rho_1$ is an element of (1.5.1), which implies that (1.5.1) contains all integer powers of $\rho_1$. In fact, one can verify that every element of (1.5.1) is an integer power of $\rho_1$ in this case.

Suppose that $|\cdot|$ is an ultrametric absolute value function on $k$. If $x, y \in k$ satisfy

$$|x - y| < |y|, \quad (1.5.4)$$

then

$$|x| = |y|. \quad (1.5.5)$$

More precisely,

$$|x| \leq \max(|x - y|, |y|) = |y|, \quad (1.5.6)$$

by the ultrametric version of the triangle inequality. Similarly,

$$|y| \leq \max(|x - y|, |x|), \quad (1.5.7)$$

which implies that $|y| \leq |x|$ in this situation.

Let $k_0$ be a subfield of $k$ that is dense with respect to the ultrametric associated to $|\cdot|$. The remarks in the preceding paragraph imply that

$$\{ |x| : x \in k_0, x \neq 0 \} \quad (1.5.8)$$

is the same as (1.5.1). In particular, if $k$ is not already complete with respect to the ultrametric associated to $|\cdot|$, then the nonzero values of the extension of $|\cdot|$ to the completion of $k$ is the same as (1.5.1).

1.6 $p$-adic integers

Let $k$ be a field, let $x$ be an element of $k$, and let $n$ be a nonnegative integer. Observe that

$$\left(1 - x\right) \sum_{j=0}^{n} x^j = \sum_{j=0}^{n} x^j - \sum_{j=0}^{n} x^{j+1} = \sum_{j=0}^{n} x^j - \sum_{j=1}^{n+1} x^j = 1 - x^{n+1}, \quad (1.6.1)$$

where $x^j$ is interpreted as being the multiplicative identity element 1 in $k$ when $j = 0$. If $x \neq 1$, then it follows that

$$\sum_{j=0}^{n} x^j = (1 - x^{n+1}) (1 - x)^{-1}. \quad (1.6.2)$$
1.6. \textit{P-ADIC INTEGERS}

Let $| \cdot |$ be an absolute value function on $k$, so that

\begin{equation}
| \sum_{j=0}^{n} x^j (1-x)^{-1} | = |x^{n+1} (1-x)^{-1}| = |x|^{n+1} |1-x|^{-1}.
\end{equation}

If $|x| < 1$, then we get that

\begin{equation}
\lim_{n \to \infty} | \sum_{j=0}^{n} x^j (1-x)^{-1} | = 0.
\end{equation}

Let $p$ be a prime number, and let $y$ be an integer. Thus $x = py$ satisfies

\begin{equation}
|x|_p = p^{-1} |y|_p \leq p^{-1} < 1,
\end{equation}

where $| \cdot |_p$ is the $p$-adic absolute value, as before. It follows that

\begin{equation}
\lim_{n \to \infty} \left| \sum_{j=0}^{n} p^j y^j (1-py)^{-1} \right|_p = 0,
\end{equation}

as in (1.6.4). Note that $\sum_{j=0}^{n} p^j y^j$ is an integer for each nonnegative integer $n$.

Suppose that $z \in \mathbb{Q}$ satisfies $|z|_p \leq 1$. This means that $z$ can be expressed as $a/b$, where $a$ and $b$ are integers, $b \neq 0$, and $b$ is not a multiple of $p$. Hence there is an integer $c$ such that $bc \equiv 1 \mod p$, because the integers modulo $p$ form a field. Thus $z$ can be expressed as

\begin{equation}
z = (a/c)/(b/c) = a/c(1-p y)^{-1},
\end{equation}

where $y$ is an integer. This implies that $z$ can be approximated by integers with respect to the $p$-adic metric, because of the analogous statement for $(1-p y)^{-1}$, as in the preceding paragraph.

Put

\begin{equation}
\mathbb{Z}_p = \{ x \in \mathbb{Q}_p : |x|_p \leq 1 \},
\end{equation}

which is the set of \textit{p-adic integers}. Of course, the set $\mathbb{Z}$ of integers is contained in $\mathbb{Z}_p$, by the definition of the $p$-adic absolute value. This implies that the closure of $\mathbb{Z}$ in $\mathbb{Q}_p$ with respect to the $p$-adic metric is contained in $\mathbb{Z}_p$, because $\mathbb{Z}_p$ is a closed set in $\mathbb{Q}_p$. Conversely, let $x \in \mathbb{Z}_p$ be given, and let us check that $x$ is in the closure of $\mathbb{Z}$ in $\mathbb{Q}_p$. Of course, $\mathbb{Q}$ is dense in $\mathbb{Q}_p$, by construction. Thus $x$ can be approximated by $z \in \mathbb{Q}$ with respect to the $p$-adic metric. In particular, if $|x-z|_p \leq 1$, then $|z|_p \leq 1$, by the ultrametric version of the triangle inequality. This means that $x$ can be approximated by $z \in \mathbb{Q}$ with $|z|_p \leq 1$ with respect to the $p$-adic metric. If $z \in \mathbb{Q}$ and $|z|_p \leq 1$, then $z$ can be approximated by integers with respect to the $p$-adic metric, as in the previous paragraph. This implies that $x$ can be approximated by integers with respect to the $p$-adic metric, as desired.
1.7 Residue fields

Let $k$ be a field, and suppose that $| \cdot |$ is an ultrametric absolute value function on $k$. It is easy to see that the open ball $B(0, r)$ in $k$ centered at 0 with radius $r > 0$ with respect to the ultrametric associated to $| \cdot |$ is a subgroup of $k$ as a commutative group with respect to addition. Similarly, the closed ball $\overline{B}(0, r)$ in $k$ centered at 0 with radius $r \geq 0$ is a subgroup of $k$ with respect to addition. The closed unit ball $B(0, 1)$ is a subring of $k$, which contains the multiplicative identity element 1 in $k$ in particular. Note that $B(0, r)$ is an ideal in $B(0, 1)$ when $0 < r \leq 1$, and that $\overline{B}(0, r)$ is an ideal in $\overline{B}(0, 1)$ when $0 \leq r \leq 1$.

Thus the quotient

$$\overline{B}(0, 1)/B(0, r)$$

(1.7.1)

can be defined as a commutative ring when $0 < r \leq 1$, and

$$\overline{B}(0, 1)/\overline{B}(0, r)$$

(1.7.2)

can be defined as a commutative ring when $0 \leq r \leq 1$. One can check that

$$\overline{B}(0, 1)/B(0, 1)$$

(1.7.3)

is a field, which is the residue field associated to $| \cdot |$ on $k$. More precisely, a nonzero element of (1.7.3) comes from an element $x$ of $\overline{B}(0, 1)$ that is not in $B(0, 1)$. This means that $|x| = 1$, so that $1/x$ is an element of $\overline{B}(0, 1)$ too. The element of (1.7.3) corresponding to $1/x$ is the inverse of the given element of (1.7.3), as desired.

If $| \cdot |$ is the trivial absolute value function on $k$, then $\overline{B}(0, 1) = k$, $B(0, 1) = \{0\}$, and the residue field (1.7.3) reduces to $k$ itself. If $k$ has characteristic $p > 0$, and $| \cdot |$ is any ultrametric absolute value function on $k$, then it is easy to see that the residue field (1.7.3) has characteristic $p$ as well.

Let $k$ be any field with an ultrametric absolute value function $| \cdot |$ again, and let $k_0$ be a subfield of $k$. The restriction of $| \cdot |$ to $k_0$ defines an ultrametric absolute value function on $k_0$, and there is a natural embedding of the residue field associated to $k_0$ into the residue field associated to $k$. If $k_0$ is dense in $k$ with respect to the ultrametric associated to $| \cdot |$, then one can check that the embedding of the residue field associated to $k_0$ into the residue field associated to $k$ is surjective, so that the residue fields are isomorphic. In particular, if $k$ is not complete with respect to the ultrametric associated to $| \cdot |$, then the residue field associated to the completion of $k$ is isomorphic to the residue field associated to $k$.

Let $p$ be a prime number, and consider $k = \mathbb{Q}_p$ with the $p$-adic absolute value. In this case,

$$\overline{B}(0, p^{-j}) = p^j \mathbb{Z}_p$$

(1.7.4)

for every $j \in \mathbb{Z}$, where $p^j \mathbb{Z}_p$ is the set of $p^j x$, $x \in \mathbb{Z}_p$. As before, $\mathbb{Z}_p$ is a subring of $\mathbb{Q}_p$, $p^j \mathbb{Z}_p$ is an ideal in $\mathbb{Z}_p$ for each nonnegative integer $j$, and hence the quotient

$$\overline{B}(0, 1)/\overline{B}(0, p^{-j}) = \mathbb{Z}_p/(p^j \mathbb{Z}_p)$$

(1.7.5)
1.8 Norms and ultranorms

Let $k$ be a field with an absolute value function $|\cdot |$, and let $V$ be a vector space over $k$. A nonnegative real-valued function $N$ on $V$ is said to be a seminorm on $V$ with respect to $|\cdot |$ if it satisfies the following two conditions. First,

\begin{equation}
N(tv) = |t| N(v)
\end{equation}

for every $t \in k$ and $v \in V$. Second,

\begin{equation}
N(v + w) \leq N(v) + N(w)
\end{equation}

for every $v, w \in V$. Note that (1.8.1) implies that $N(0) = 0$, by taking $t = 0$. If we also have that

\begin{equation}
N(v) > 0
\end{equation}

for every $v \in V$ with $v \neq 0$, then $N$ is said to be a norm on $V$ with respect to $|\cdot |$. In particular, $k$ may be considered as a one-dimensional vector space over itself, and $|\cdot |$ may be considered as a norm on $k$ with respect to itself.

A nonnegative real-valued function $N$ on $V$ is said to be a semi-ultranorm on $V$ with respect to $|\cdot |$ on $k$ if it satisfies (1.8.1) and

\begin{equation}
N(v + w) \leq \max(N(v), N(w))
\end{equation}

for every $v, w \in V$. If $N$ also satisfies (1.8.3), then $N$ is said to be an ultranorm on $V$ with respect to $|\cdot |$. Of course, (1.8.4) implies (1.8.2), so that semi-ultranorms and ultranorms are seminorms and ultranorms, respectively. If $N$ is a semi-ultranorm on $V$ with respect to $|\cdot |$ on $k$, and if $N(v) > 0$ for some $v \in V$, then one can check that $|\cdot |$ is an ultrametric absolute value function on $k$. If $|\cdot |$ is an ultrametric absolute value function on $k$, then $|\cdot |$ may be considered as an ultranorm on $k$ as a one-dimensional vector space over itself.
CHAPTER 1. ABSOLUTE VALUES AND NORMS

Let $| \cdot |$ be any absolute value function on $k$ again. If $N$ is a seminorm on $V$ with respect to $| \cdot |$, then

$$d(v, w) = d_N(v, w) = N(v - w)$$

(1.8.5)

defines a semimetric on $V$. If $N$ is a norm on $V$, then (1.8.5) is a metric on $V$. If $N$ is a semi-ultranorm on $V$, then (1.8.5) is a semi-ultrametric on $V$. Thus (1.8.5) is an ultrametric on $V$ when $N$ is an ultranorm on $V$.

Suppose for the moment that $| \cdot |$ is the trivial absolute value function on $k$. The trivial ultranorm is defined on $V$ by putting $N(v)$ equal to 1 when $v \neq 0$, and equal to 0 when $v = 0$. It is easy to see that this defines an ultranorm on $V$, for which the corresponding ultrametric is the discrete metric.

Let $| \cdot |$ be any absolute value function on $k$, and let $a$ be a positive real number with $a \leq 1$. Remember that $| \cdot |^a$ defines an absolute value function on $k$ too, as in Section 1.3. If $N$ is a seminorm on $V$ with respect to $| \cdot |$ on $k$, then $N(v)^a$ is a seminorm on $V$ with respect to $| \cdot |^a$ on $k$. This uses (1.3.1) to get the triangle inequality for $N(v)^a$ from the one for $| \cdot |^a$ on $k$.

If $| \cdot |$ is an ultrametric absolute value function on $k$, then $| \cdot |^a$ is an ultrametric absolute value function on $k$ for every $a > 0$, as in Section 1.3. In this case, if $N$ is a semi-ultranorm on $V$ with respect to $| \cdot |$ on $k$, then $N(v)^a$ is a semi-ultranorm on $V$ with respect to $| \cdot |^a$ for every $a > 0$. If $N$ is an ultranorm on $V$ with respect to $| \cdot |$ on $k$, then $N(v)^a$ is an ultranorm on $V$ with respect to $| \cdot |^a$ on $k$ for every $a > 0$.

Let $N$ be a semi-ultranorm on $V$ with respect to $| \cdot |$ on $k$. If $v, w \in V$ satisfy

$$N(v - w) < N(w),$$

(1.8.6)

then

$$N(v) = N(w).$$

(1.8.7)

This is analogous to the corresponding statement for ultrametric absolute value functions mentioned in Section 1.5. If $N(v - w) \leq N(w)$, then

$$N(v) \leq \max(N(v - w), N(w)) = N(w),$$

(1.8.8)

by the semi-ultranorm version of the triangle inequality. We also have that

$$N(w) \leq \max(N(v - w), N(v)),$$

(1.8.9)

which implies that $N(w) \leq N(v)$ when (1.8.6) holds.

Suppose that $N$ is a norm on $V$ with respect to an absolute value function $| \cdot |$ on $k$. If $V$ is complete with respect to the metric associated to $N$, then $V$ is said to be a Banach space with respect to $N$. Otherwise, one can pass to a completion of $V$. The vector space operations on $V$ can be extended to the completion, so that the completion becomes a vector space over $k$. The extension of $N$ to the completion corresponds to the distance to 0 on the completion, and defines a norm on the completion. If $N$ is an ultranorm on $V$, then the extension of $N$
1.9. BOUNDED LINEAR MAPPINGS

Let $k$ be a field with an absolute value function $| \cdot |$, and let $V$, $W$ be vector spaces over $k$. Also let $N_V$, $N_W$ be seminorms on $V$, $W$, respectively, with respect to $| \cdot |$ on $k$. A linear mapping $T$ from $V$ into $W$ is said to be \textit{bounded} with respect to $N_V$ and $N_W$ if there is a nonnegative real number $C$ such that

$$N_W(T(v)) \leq C N_V(v)$$

for every $v \in V$. This implies that

$$N_W(T(v) - T(v')) = N_W(T(v - v')) \leq C N_V(v - v')$$

for every $v, v' \in V$, and in particular that $T$ is continuous with respect to the seminorms associated to $N_V$ and $N_W$ on $V$ and $W$, respectively. Conversely, if a linear mapping $T$ from $V$ into $W$ is continuous at 0 with respect to these seminorms, and if $| \cdot |$ is not the trivial absolute value function on $k$, then one can check that $T$ is bounded with respect to $N_V$ and $N_W$.

Let $\mathcal{B}(V, W)$ be the space of bounded linear mappings from $V$ into $W$, with respect to $N_V$ and $N_W$. If $T \in \mathcal{B}(V, W)$, then put

$$\|T\|_{op} = \|T\|_{op,VW} = \inf \{ C \geq 0 : (1.9.1) \text{ holds} \},$$

where more precisely the infimum is taken over all nonnegative real numbers $C$ such that (1.9.1) holds for every $v \in V$. Note that the infimum is automatically attained in this situation, which is to say that (1.9.1) holds with $C = \|T\|_{op}$. One can verify that $\mathcal{B}(V, W)$ is a vector space over $k$ with respect to pointwise addition and scalar multiplication of mappings from $V$ into $W$, and that (1.9.3) defines a seminorm on $\mathcal{B}(V, W)$ with respect to $| \cdot |$ on $k$. If $N_V$ is a norm on $W$, then (1.9.3) is a norm on $\mathcal{B}(V, W)$. If $N_W$ is a semi-ultranorm on $W$, then (1.9.3) is a semi-ultranorm on $\mathcal{B}(V, W)$. In particular, if $N_W$ is an ultranorm on $W$, then (1.9.3) is an ultranorm on $\mathcal{B}(V, W)$.

Let $Z$ be another vector space over $k$, and let $N_Z$ be a seminorm on $Z$ with respect to $| \cdot |$ on $k$. Suppose that $T_1$ is a bounded linear mapping from $V$ into $W$ with respect to $N_V$ and $N_W$, and that $T_2$ is a bounded linear mapping from $W$ into $Z$ with respect to $N_W$ and $N_Z$. If $v \in V$, then

$$N_Z(T_2(T_1(v))) \leq \|T_2\|_{op,WZ} N_W(T_1(v)) \leq \|T_1\|_{op,VW} \|T_2\|_{op,WZ} N_V(v),$$

to the completion of $V$ is an ultranorm as well. The completion of $V$ is unique, up to isometric isomorphic equivalence.

If $N$ is a seminorm on $V$ with respect to $| \cdot |$ on $k$, and if $V_0$ is a linear subspace of $V$, then the restriction of $N$ to $V_0$ is a seminorm on $V$ with respect to $| \cdot |$ on $k$. If $V_0$ is dense in $V$ with respect to the semimetric associated to $N$, and if the restriction of $N$ to $V_0$ is a semi-ultranorm on $V_0$, then it is easy to see that $N$ is a semi-ultranorm on $V$. In particular, if $N$ is an ultranorm on $V$, and if $V$ is not already complete with respect to the ultrametric associated to $N$, then the extension of $N$ to the completion of $V$ is an ultranorm.
where the subscripts indicate the spaces involved in the corresponding operator seminorm. This implies that the composition \( T_2 \circ T_1 \) is bounded as a linear mapping from \( V \) into \( Z \), with

\[
\| T_2 \circ T_1 \|_{\text{op};VZ} \leq \| T_1 \|_{\text{op};VW} \| T_2 \|_{\text{op};WZ}.
\]

Let us suppose from now on in this section that \( N_W \) is a norm on \( W \), and that \( W \) is complete with respect to the metric associated to \( N_W \). Under these conditions, one can check that \( BL(V,W) \) is complete with respect to the operator norm (1.9.3), using standard arguments. More precisely, if \( \{ T_j \}_{j=1}^\infty \) is a Cauchy sequence in \( BL(V,W) \) with respect to the metric associated to the operator norm, then \( \{ T_j(v) \}_{j=1}^\infty \) is a Cauchy sequence in \( W \) with respect to the metric associated to \( N_W \) for every \( v \in V \). This implies that \( \{ T_j(v) \}_{j=1}^\infty \) converges to a unique element \( T(v) \) of \( W \) with respect to the metric associated to \( N_W \), because \( W \) is supposed to be complete with respect to this metric. It is easy to see that \( T \) defines a linear mapping from \( V \) into \( W \), because \( T_j \) is linear for each \( j \). The Cauchy condition for \( \{ T_j \}_{j=1}^\infty \) with respect to the metric associated to the operator norm implies that the operator norms of the \( T_j \)'s are bounded, which can be used to get that \( T \) is a bounded linear mapping. One can use the Cauchy condition for \( \{ T_j \}_{j=1}^\infty \) again to obtain that this sequence converges to \( T \) with respect to the metric associated to the operator norm, as desired.

Suppose for convenience that \( N_V \) is a norm on \( V \), although this is not really needed. Let \( V_0 \) be a linear subspace of \( V \) that is dense in \( V \) with respect to the metric associated to \( N_V \), and let \( T_0 \) be a bounded linear mapping from \( V_0 \) into \( W \), with respect to the restriction of \( N_V \) to \( V_0 \). Note that \( T_0 \) is uniformly continuous with respect to the metric on \( V_0 \) associated to the restriction of \( N_V \) to \( V_0 \), and the metric on \( W \) associated to \( N_W \), as in (1.9.2). It follows that there is a unique extension of \( T_0 \) to a uniformly continuous mapping from \( V \) into \( W \), with respect to the metrics associated to \( N_V \) and \( N_W \), respectively, as mentioned at the beginning of Section 1.4. One can check that this extension is a bounded linear mapping from \( V \) into \( W \) with respect to \( N_V \) and \( N_W \), with the same operator norm as \( T_0 \) has on \( V_0 \).

### 1.10 Some norms on \( \mathbb{k}^n \)

Let \( \mathbb{k} \) be a field, and let \( n \) be a positive integer. The space \( \mathbb{k}^n \) of \( n \)-tuples \( v = (v_1, \ldots, v_n) \) of elements of \( \mathbb{k} \) is a vector space over \( \mathbb{k} \) with respect to coordinatewise addition and scalar multiplication. Let \(| \cdot |\) be an absolute value function on \( \mathbb{k} \). It is easy to see that

\[
\| v \|_1 = \sum_{j=1}^n |v_j|
\]

and

\[
\| v \|_\infty = \max_{1 \leq j \leq n} |v_j|
\]
are norms on $k^n$ with respect to $| \cdot |$ on $k$. If $| \cdot |$ is an ultrametric absolute value function on $k$, then (1.10.2) is an ultranorm on $k^n$.

Observe that
\[\|v\|_\infty \leq \|v\|_1 \leq n \|v\|_\infty\]
for every $v \in k^n$. Let
\[d_1(v, w) = \|v - w\|_1\]
and
\[d_\infty(v, w) = \|v - w\|_\infty\]
be the metrics on $k^n$ associated to (1.10.1) and (1.10.2), respectively. Thus
\[d_\infty(v, w) \leq d_1(v, w) \leq n d_\infty(v, w)\]
for every $v, w \in k^n$, by (1.10.3). In particular, this implies that (1.10.4) and (1.10.5) determine the same topology on $k^n$, corresponding to the topology determined on $k$ by the metric associated to $| \cdot |$.

The standard basis vectors $e_1, \ldots, e_n$ in $k^n$ are defined as usual by taking the $j$th coordinate of $e_l$ to be equal to 1 when $j = l$ and to 0 when $j \neq l$, where $1 \leq j, l \leq n$. Thus
\[v = \sum_{l=1}^{n} v_l e_l\]
for every $v \in k^n$. Let $W$ be a vector space over $k$, and let $N_W$ be a seminorm on $W$ with respect to $| \cdot |$ on $k$. If $T$ is a linear mapping from $k^n$ into $W$, then
\[T(v) = T\left(\sum_{l=1}^{n} v_l e_l\right) = \sum_{l=1}^{n} v_l T(e_l)\]
for every $v \in k^n$, and hence
\[N_W(T(v)) \leq \sum_{l=1}^{n} |v_l| N_W(T(e_l)).\]
In particular,
\[N_W(T(v)) \leq \left(\max_{1 \leq l \leq n} N_W(T(e_l))\right) \|v\|_1\]
for every $v \in k^n$. This means that $T$ is bounded as a linear mapping from $k^n$ equipped with $\|v\|_1$ into $W$, with operator seminorm less than or equal to
\[\max_{1 \leq l \leq n} N_W(T(e_l)).\]
In fact, the operator seminorm of $T$ is equal to (1.10.11) in this situation, because the operator seminorm of $T$ is automatically greater than or equal to $N_W(T(e_l))$ for each $l = 1, \ldots, n$, since $\|e_l\|_1 = 1$. Similarly,
\[N_W(T(v)) \leq \left(\sum_{l=1}^{n} N_W(T(e_l))\right) \|v\|_\infty\]
for every $v \in k^n$, by (1.10.9). This implies that $T$ is bounded as a linear mapping from $k^n$ equipped with $\|v\|_\infty$ into $W$, with operator seminorm less than or equal to
\begin{equation}
\sum_{l=1}^{n} N_W(T(e_l)).
\end{equation}
Note that the operator seminorm of $T$ is greater than or equal to (1.10.11), because $\|e_l\|_\infty = 1$ for each $l = 1, \ldots, n$.

Suppose now that $| \cdot |$ is an ultrametric absolute value function on $k$, and that $N_W$ is a semi-ultranorm on $W$ with respect to $| \cdot |$ on $k$. Using (1.10.8), we get that
\begin{equation}
N_W(T(v)) \leq \max_{1 \leq l \leq n} (|v| \cdot N_W(T(e_l))) \leq \left( \max_{1 \leq l \leq n} N_W(T(e_l)) \right) \|v\|_\infty
\end{equation}
for every $v \in k^n$. This implies that $T$ is bounded as a linear mapping from $k^n$ equipped with $\|v\|_\infty$ into $W$, with operator seminorm less than or equal to (1.10.11). The operator seminorm of $T$ is also greater than or equal to (1.10.11), as in the preceding paragraph. Hence the operator seminorm of $T$ with respect to $\|v\|_\infty$ on $k^n$ is equal to (1.10.11) in this case.

### 1.11 Inner products

Suppose for the moment that $V$ and $W$ are vector spaces over the field $\mathbb{C}$ of complex numbers, so that $V$ and $W$ may be considered as vector spaces over $\mathbb{R}$ as well. Let us say that a mapping $T$ from $V$ into $W$ is real-linear if $T$ is linear as a mapping from $V$ into $W$ as vector spaces over $\mathbb{R}$, and that $T$ is complex-linear if $T$ is linear as a mapping from $V$ into $W$ as vector spaces over $\mathbb{C}$. Thus a complex-linear mapping $T$ from $V$ into $W$ is the same as a real-linear mapping that also satisfies
\begin{equation}
T(i v) = iT(v)
\end{equation}
for every $v \in V$. A real-linear mapping $T$ from $V$ into $W$ is said to be conjugate-linear if
\begin{equation}
T(i v) = -iT(v)
\end{equation}
for every $v \in V$. This implies that
\begin{equation}
T(a v) = \overline{a}T(v)
\end{equation}
for every $a \in \mathbb{C}$ and $v \in V$, where $\overline{a}$ is the usual complex-conjugate of $a$.

Suppose from now on in this section that $k = \mathbb{R}$ or $\mathbb{C}$, with the standard absolute value function. Let $V$ be a vector space over $k$, and let $\langle v, w \rangle$ be a $k$-valued function defined for $v, w \in V$. If the following three conditions are satisfied, then $\langle v, w \rangle$ is said to be an inner product on $V$. The first condition is that $\langle v, w \rangle$ be linear in $v$ for each $w \in V$. The second condition is that
\begin{equation}
\langle w, v \rangle = \langle v, w \rangle
\end{equation}
for every \( v, w \in V \) in the real case, and that
\[
\langle v, w \rangle = \overline{\langle w, v \rangle}
\]
for every \( v, w \in V \) in the complex case. Note that \( \langle v, w \rangle \) is linear in \( w \) for each \( v \in V \) in the real case, and conjugate-linear in \( w \) for each \( v \in V \) in the complex case. In the complex case, we also get that \( \langle v, v \rangle \) is a real number for every \( v \in V \), by (1.11.5). The third condition is that
\[
\langle v, v \rangle > 0
\]
for every \( v \in V \) with \( v \neq 0 \). Of course, \( \langle v, w \rangle = 0 \) when either \( v = 0 \) or \( w = 0 \), by the first two conditions. If \( \langle v, w \rangle \) is an inner product on \( V \), then we put
\[
\|v\| = \langle v, v \rangle^{1/2}
\]
for every \( v \in V \), using the nonnegative square root on the right side. It is well known that
\[
|\langle v, w \rangle| \leq \|v\| \|w\|
\]
for every \( v, w \in V \), which is the Cauchy–Schwarz inequality. Using this, one can show that \( \| \cdot \| \) defines a norm on \( V \). If \( V \) is complete with respect to the metric associated to \( \| \cdot \| \), then \( V \) is said to be a Hilbert space with respect to \( \langle v, w \rangle \). Otherwise, one can pass to a completion, as usual.

Let \( n \) be a positive integer. The standard inner product on \( \mathbb{R}^n \) is given by
\[
\langle v, w \rangle = \langle v, w \rangle_{\mathbb{R}^n} = \sum_{j=1}^{n} v_j w_j.
\]
Similarly, the standard inner product on \( \mathbb{C}^n \) is given by
\[
\langle v, w \rangle = \langle v, w \rangle_{\mathbb{C}^n} = \sum_{j=1}^{n} v_j \overline{w_j}.
\]
In both cases, the corresponding norm is given by
\[
\|v\| = \|v\|_2 = \left( \sum_{j=1}^{n} |v_j|^2 \right)^{1/2}.
\]
It is easy to see that
\[
\|v\|_{\infty} \leq \|v\|_2 \leq n^{1/2} \|v\|_{\infty}
\]
for every \( v \in \mathbb{R}^n \) or \( \mathbb{C}^n \), where \( \|v\|_{\infty} \) is as in (1.10.2). One can also check that
\[
\|v\|_2 \leq \|v\|_1 \leq n^{1/2} \|v\|_2
\]
for every \( v \in \mathbb{R}^n \) or \( \mathbb{C}^n \), where \( \|v\|_1 \) is as in (1.10.1). More precisely, the first inequality in (1.11.13) can be verified using the first inequality in (1.10.3), and
CHAPTER 1. ABSOLUTE VALUES AND NORMS

the second inequality in (1.11.13) can be obtained from the Cauchy–Schwarz inequality.

Let \((V, \langle \cdot, \cdot \rangle_V)\) and \((W, \langle \cdot, \cdot \rangle_W)\) be Hilbert spaces, both real or both complex, and let \(\| \cdot \|_V\) and \(\| \cdot \|_W\) be the corresponding norms on \(V\) and \(W\), respectively. Also let \(T\) be a bounded linear mapping from \(V\) into \(W\), with respect to \(\| \cdot \|_V\) and \(\| \cdot \|_W\). It is well known that there is a unique bounded linear mapping \(T^*\) from \(W\) into \(V\) such that

\[
(T(v), w)_W = \langle v, T^*(w) \rangle_V
\]

(1.11.14)

for every \(v \in V\) and \(w \in W\). This mapping \(T^*\) is called the adjoint of \(T\). The adjoint \((T^*)^*\) of \(T^*\) can be defined as a bounded linear mapping from \(V\) into \(W\) in the same way, and is equal to \(T\). It is not difficult to show that

\[
\|T^*\|_{op,WV} = \|T\|_{op,VW}.
\]

(1.11.15)

Note that \(T \mapsto T^*\) is a linear mapping from \(\mathcal{B}(V,W)\) into \(\mathcal{B}(W,V)\) in the real case, and that this mapping is conjugate-linear in the complex case.

Let \((Z, \langle \cdot, \cdot \rangle_Z)\) be another Hilbert space, which is real when \(V\) and \(W\) are real, and complex when \(V\) and \(W\) are complex, and let \(\| \cdot \|_Z\) be the corresponding norm on \(Z\). If \(T_1\) is a bounded linear mapping from \(V\) into \(W\), and \(T_2\) is a bounded linear mapping from \(W\) into \(Z\), then their composition \(T_2 \circ T_1\) is a bounded linear mapping from \(V\) into \(Z\), as before. It is easy to see that

\[
(T_2 \circ T_1)^* = T_1^* \circ T_2^*
\]

(1.11.16)
as bounded linear mappings from \(Z\) into \(V\).

1.12 Infinite series

Let \(k\) be a field with an absolute value function \(|\cdot|\), and let \(V\) be a vector space over \(k\) with a norm \(N\) with respect to \(|\cdot|\) on \(k\). An infinite series \(\sum_{j=1}^{\infty} v_j\) with terms in \(V\) is said to converge in \(V\) if the corresponding sequence of partial sums \(\sum_{j=1}^{n} v_j\) converges to an element of \(V\) with respect to the metric associated to \(N\). In this case, the value of the sum \(\sum_{j=1}^{\infty} v_j\) is defined to be the limit of the sequence of partial sums. If \(\sum_{j=1}^{\infty} v_j\) converges in \(V\) and \(t \in k\), then it is easy to see that \(\sum_{j=1}^{\infty} tv_j\) converges in \(V\) too, with

\[
\sum_{j=1}^{\infty} tv_j = t \sum_{j=1}^{\infty} v_j.
\]

(1.12.1)

Similarly, if \(\sum_{j=1}^{\infty} w_j\) is another convergent series in \(V\), then \(\sum_{j=1}^{\infty} (v_j + w_j)\) converges in \(V\) as well, with

\[
\sum_{j=1}^{\infty} (v_j + w_j) = \sum_{j=1}^{\infty} v_j + \sum_{j=1}^{\infty} w_j.
\]

(1.12.2)
A necessary condition for the convergence of an infinite series $\sum_{j=1}^{\infty} v_j$ with terms in $V$ is that the corresponding sequence of partial sums be a Cauchy sequence with respect to the metric associated to $N$. This happens if and only if for every $\epsilon > 0$ there is a positive integer $L$ such that

$$N\left(\sum_{j=l}^{n} v_j\right) < \epsilon$$

(1.12.3)

for all $l, n \in \mathbb{Z}_+$ with $n \geq l \geq L$. In particular, this implies that

$$\lim_{j \to \infty} N(v_j) = 0,$$

(1.12.4)

by taking $l = n$ in (1.12.3). Of course, if $V$ is complete with respect to the metric associated to $N$, then the Cauchy condition (1.12.3) implies that $\sum_{j=1}^{\infty} v_j$ converges in $V$.

If $\sum_{j=1}^{\infty} N(v_j)$ converges as an infinite series of nonnegative real numbers, then $\sum_{j=1}^{\infty} v_j$ is said to converge absolutely with respect to $N$. Observe that

$$N\left(\sum_{j=l}^{n} v_j\right) \leq \sum_{j=l}^{n} N(v_j)$$

(1.12.5)

for every $l, n \in \mathbb{Z}_+$ with $n \geq l$, by the triangle inequality for $N$. If $\sum_{j=1}^{\infty} v_j$ converges absolutely with respect to $N$, then it is easy to see that the Cauchy condition (1.12.3) holds, using (1.12.5). If $V$ is complete with respect to the metric associated to $N$, then it follows that $\sum_{j=1}^{\infty} v_j$ converges in $V$. In this case, we also have that

$$N\left(\sum_{j=1}^{\infty} v_j\right) \leq \sum_{j=1}^{\infty} N(v_j).$$

(1.12.6)

Suppose for the moment that $N$ is an ultranorm on $V$ with respect to $\| \cdot \|$ on $k$, so that

$$N\left(\sum_{j=l}^{n} v_j\right) \leq \max_{l \leq j \leq n} N(v_j)$$

(1.12.7)

for every $n \geq l \geq 1$. If (1.12.4) holds, then it follows that the Cauchy condition (1.12.3) holds too. If $V$ is complete with respect to the ultrametric associated to $N$, then we get that $\sum_{j=1}^{\infty} v_j$ converges in $V$. Note that

$$N\left(\sum_{j=1}^{\infty} v_j\right) \leq \max_{j \geq 1} N(v_j)$$

(1.12.8)

in this situation. More precisely, the maximum on the right side of (1.12.8) is attained, because of (1.12.4).
Let us now take \( k = \mathbb{R} \) or \( \mathbb{C} \), with the standard absolute value function. Let \( (V, \langle \cdot, \cdot \rangle) \) be a real or complex inner product space, and let \( \| \cdot \| \) be the corresponding norm on \( V \), as in the previous section. Suppose that \( \sum_{j=1}^{\infty} v_j \) is an infinite series of pairwise-orthogonal vectors in \( V \), so that
\[
\langle v_j, v_l \rangle = 0 \quad \text{when} \ j \neq l.
\]
This implies that
\[
\sum_{j=l}^{n} v_j^2 = \sum_{j=l}^{n} \| v_j \|^2 \quad \text{for every} \ n \geq l \geq 1.
\]
If \( \sum_{j=1}^{\infty} \| v_j \|^2 \) converges as an infinite series of nonnegative real numbers, then the Cauchy condition (1.12.3) holds, with \( N = \| \cdot \| \). Hence \( \sum_{j=1}^{\infty} v_j \) converges in \( V \) when \( V \) is a Hilbert space, in which case we have that
\[
\sum_{j=1}^{\infty} v_j^2 = \sum_{j=1}^{\infty} \| v_j \|^2.
\]
Conversely, if the Cauchy condition (1.12.3) holds, then \( \sum_{j=1}^{\infty} \| v_j \|^2 \) converges, because the partial sums are bounded.

### 1.13 Bounded bilinear mappings

Let \( k \) be a field, and let \( V, W, \) and \( Z \) be vector spaces over \( k \). A mapping \( b \) from \( V \times W \) into \( Z \) is said to be bilinear if \( b(v, w) \) is linear in \( v \) for each \( w \in W \), and linear in \( w \) for each \( v \in V \). Let \( | \cdot | \) be an absolute value function on \( k \), and let \( N_V, N_W, \) and \( N_Z \) be seminorms on \( V, W, \) and \( Z \), respectively, with respect to \( \| \cdot \| \) on \( k \). If there is a nonnegative real number \( C \) such that
\[
N_Z(b(v, w)) \leq C N_V(v) N_W(w)
\]
for every \( v \in V \) and \( w \in W \), then \( b \) is said to be bounded as a bilinear mapping from \( V \times W \) into \( Z \).

Let \( b \) be a bilinear mapping from \( V \times W \) into \( Z \) that satisfies (1.13.1), and let \( v, v' \in V \) and \( w, w' \in W \) be given. Observe that
\[
b(v, w) - b(v', w) = b(v-v', w) + b(v', w-w'),
\]
so that
\[
N_Z(b(v, w) - b(v', w')) \leq N_Z(b(v-v', w)) + N_Z(b(v', w-w')) \leq C \| v - v' \|_V \| w \| + C \| v' \|_V \| w - w' \|_W.
\]
One can use this to check that \( b \) is continuous with respect to the semimetrics associated to \( N_V, N_W, N_Z \) and the corresponding product topology on \( V \times W \).
1.13. BOUNDED BILINEAR MAPPINGS

Conversely, if a bilinear mapping $b$ from $V \times W$ into $Z$ is continuous at $(0,0)$ in $V \times W$ with respect to these semimetrics and the corresponding product topology on $V \times W$, and if $| \cdot |$ is not the trivial absolute value function on $k$, then one can verify that $b$ is bounded as a bilinear mapping.

Suppose for the moment that $N_V$, $N_W$, and $N_Z$ are norms on $V$, $W$, and $Z$, respectively, and let $V_0$, $W_0$ be dense linear subspaces of $V$ and $W$ with respect to the metrics associated to $N_V$ and $N_W$. Let $b_0$ be a bounded bilinear mapping from $V_0 \times W_0$ into $Z$, using the restrictions of $N_V$ and $N_W$ to $V_0$ and $W_0$, respectively. If $Z$ is complete with respect to the metric associated to $N_Z$, then there is a unique extension of $b_0$ to a bounded bilinear mapping from $V \times W$ into $Z$. More precisely, for each $w \in W_0$, one can first extend $b_0(v,w)$ to a bounded linear mapping from $V$ into $Z$, as a function of $v$. This defines a bounded bilinear mapping from $V \times W_0$ into $Z$, which can be extended to a bounded bilinear mapping from $V \times W$ into $Z$ in the same way.

Suppose now that $V = k^{n_V}$ and $W = k^{n_W}$ for some positive integers $n_V$ and $n_W$, and let $N_Z$ be any seminorm on $Z$ again. Also let $e_1^{n_V}, \ldots, e_{n_V}^{n_V}$ and $e_1^{n_W}, \ldots, e_{n_W}^{n_W}$ be the standard basis vectors in $k^{n_V}$ and $k^{n_W}$, respectively, and let $b$ be a bilinear mapping from $k^{n_V} \times k^{n_W}$ into $Z$. Observe that

\begin{equation}
(1.13.4) \quad b(v, w) = \sum_{j=1}^{n_V} \sum_{l=1}^{n_W} v_j w_l b(e_j^{n_V}, e_l^{n_W})
\end{equation}

for every $v \in V$ and $w \in W$, so that

\begin{equation}
(1.13.5) \quad N_Z(b(v, w)) \leq \sum_{j=1}^{n_V} \sum_{l=1}^{n_W} |v_j| |w_l| N_Z(b(e_j^{n_V}, e_l^{n_W})).
\end{equation}

If we take $N_V(v)$ to be

\begin{equation}
(1.13.6) \quad \|v\|_{1,n_V} = \sum_{j=1}^{n_V} |v_j|,
\end{equation}

and $N_W(w)$ to be

\begin{equation}
(1.13.7) \quad \|w\|_{1,n_W} = \sum_{l=1}^{n_W} |w_l|,
\end{equation}

as in (1.10.1), then (1.13.5) implies that (1.13.1) holds with $C$ equal to

\begin{equation}
(1.13.8) \quad \max\{N_Z(b(e_j^{n_V}, e_l^{n_W})) : 1 \leq j \leq n_V, \ 1 \leq l \leq n_W\}.
\end{equation}

Similarly, if we take $N_V(v)$ to be

\begin{equation}
(1.13.9) \quad \|v\|_{\infty,n_V} = \max_{1 \leq j \leq n_V} |v_j|,
\end{equation}

and $N_W(w)$ to be

\begin{equation}
(1.13.10) \quad \|w\|_{\infty,n_W} = \max_{1 \leq l \leq n_W} |w_l|,
\end{equation}

then (1.13.5) implies that (1.13.1) holds with $C$ equal to

\begin{equation}
(1.13.11) \quad \max\{N_Z(b(e_j^{n_V}, e_l^{n_W})) : 1 \leq j \leq n_V, \ 1 \leq l \leq n_W\}.
\end{equation}
as in (1.10.2), then (1.13.5) implies that (1.13.1) holds, with $C$ equal to

$$\sum_{j=1}^{n_V} \sum_{l=1}^{n_W} N_Z(b(e_j^{n_V}, e_l^{n_W})).$$

If $N_Z$ is a semi-ultranorm on $Z$, then we get that

$$N_Z(b(v, w)) \leq \max\{|v_j| |w_l| N(b(e_j^{n_V}, e_l^{n_W})) : 1 \leq j \leq n_V, 1 \leq l \leq n_W\}$$

for every $v \in k^{n_V}$ and $k^{n_W}$. In this case, if we take $N_V(v)$ and $N_W(w)$ to be as in (1.13.9) and (1.13.10), respectively, then (1.13.1) holds with $C$ equal to (1.13.8).

### 1.14 Minkowski functionals

Let $k$ be a field with an absolute value function $|\cdot|$, and let $V$ be a vector space over $k$. If $t \in k$ and $E \subseteq V$, then we put

$$tE = \{tv : v \in E\}.$$

Let us say that $E$ is balanced in $V$ if

$$tE \subseteq E$$

for every $t \in k$ with $|t| \leq 1$. If $|t| = 1$, then it follows that

$$tE = E,$$

by applying (1.14.2) to both $t$ and $1/t$. Note that a nonempty balanced subset of $V$ contains 0.

Let us say that a balanced set $A \subseteq V$ is absorbing if for every $v \in V$ there is a $t_1 \in k$ such that $t_1 \neq 0$ and

$$v \in t_1A.$$

This implies that

$$v \in tA$$

for every $t \in k$ such that $|t| \geq |t_1|$, because $A$ is balanced. Equivalently, this means that

$$t^{-1}v \in A$$

when $|t| \geq |t_1|$. Of course, if (1.14.4) holds with $t_1 = 0$, then $v = 0$. We also have that $0 \in A$, because $A$ is balanced and nonempty, so that (1.14.5) holds for every $t \in k$. Clearly $V$ is automatically balanced and absorbing as a subset of itself. If $|\cdot|$ is the trivial absolute value function on $k$, then $V$ is the only balanced absorbing subset of itself.
Let \( N \) be a nonnegative real-valued function on \( V \) such that
\[
N(t v) = |t| N(v)
\]
for every \( t \in k \) and \( v \in V \). Put
\[
B_N(0, r) = \{ v \in V : N(v) < r \}
\]
for every positive real number \( r \), and
\[
\overline{B}(0, r) = \{ v \in V : N(v) \leq r \}
\]
for every nonnegative real number \( r \). If \( t \in k \) and \( t \neq 0 \), then
\[
t B_N(0, r) = B_N(0, |t| r)
\]
for every \( r > 0 \), and
\[
t \overline{B}_N(0, r) = \overline{B}_N(0, |t| r)
\]
for every \( r \geq 0 \). In particular, \( B_N(0, r) \) is balanced in \( V \) for every \( r > 0 \), and \( \overline{B}_N(0, r) \) is balanced in \( V \) for every \( r \geq 0 \). If \( |\cdot| \) is not the trivial absolute value function on \( k \), then \( B_N(0, r) \) and \( \overline{B}_N(0, r) \) are absorbing in \( V \) for every \( r > 0 \).

Let us suppose from now on in this section that \( |\cdot| \) is not the trivial absolute value function on \( k \). Let \( A \) be a balanced absorbing subset of \( V \), and put
\[
N_A(v) = \inf \{ |t| : t \in k, t \neq 0, v \in t A \}
\]
for each \( v \in V \). Note that \( N_A(0) = 0 \), so that we could have included the possibility of \( t = 0 \) in the first formulation of \( N_A(v) \). Of course, \( N_A \) is a nonnegative real-valued function on \( V \), and one can check that
\[
N_A(t' v) = |t'| N_A(v)
\]
for every \( t' \in k \) and \( v \in V \). If \( v \in A \), then \( N_A(v) \leq 1 \), so that
\[
A \subseteq \overline{B}_{N_A}(0, 1).
\]
If \( v \in V \) satisfies \( N_A(v) < 1 \), then there is a \( t \in k \) such that \( |t| < 1 \) and \( v \in t A \). This implies that \( v \in A \), because \( A \) is balanced in \( V \), so that
\[
B_{N_A}(0, 1) \subseteq A.
\]
Suppose for the moment that \( |\cdot| \) is discrete on \( k \), as in Section 1.5. If \( v \in V \) and \( N_A(v) > 0 \), then it follows that the infimum in (1.14.12) is attained. If \( v \in V \) and \( N_A(v) \leq 1 \), then there is a \( t \in k \) such that \( |t| \leq 1 \) and \( v \in t A \), because the infimum is attained when \( N_A(v) = 1 \). As before, this implies that \( v \in A \), because \( A \) is balanced in \( V \). Hence
\[
A = \overline{B}_{N_A}(0, 1)
\]
in this situation.
1.15 Balanced subgroups

Let $k$ be a field with an absolute value function $| \cdot |$ again, and let $V$ be a vector space over $k$. Let us say that $E \subseteq V$ is a balanced subgroup of $V$ if $E$ is balanced as a subset of $V$, as in the previous section, and $E$ is a subgroup of $V$ as a commutative group with respect to addition. If $E \subseteq V$ is a nonempty balanced subset of $V$, then $0 \in E$ and $-E = E$. If we also have that $v + w \in E$ for every $v;w \in E$, then it follows that $E$ is a balanced subgroup of $V$.

Let $N$ be a semi-ultranorm on $V$ with respect to $| \cdot |$ on $k$. Note that $N_A(0,r)$ in (1.14.8) is the same as the open ball in $V$ centered at 0 with radius $r > 0$ with respect to the semi-ultrametric associated to $N$, and that (1.14.9) is the same as the closed ball in $V$ centered at 0 with radius $r \geq 0$ with respect to the semi-ultrametric associated to $N$. It is easy to see that $N_A(0,r)$ is a balanced subgroup in $V$ for every $r > 0$, and that $\overline{N_A}(0,r)$ is a balanced subgroup for every $r \geq 0$. More precisely, (1.15.1) holds in both cases, by the semi-ultranorm version of the triangle inequality.

Of course, linear subspaces of $V$ are balanced subgroups of $V$. If $| \cdot |$ is archimedean on $k$, then one can check that any balanced subgroup $E$ of $V$ is linear subspace of $V$. This uses the fact that for each $v \in E$ and positive integer $n$, the sum of $n$ $v$'s in $V$ is an element of $E$. If $| \cdot |$ is the trivial absolute value function on $k$, then balanced subgroups of $V$ are linear subspaces again.

If $E$ is a balanced subgroup in $V$, and if the linear span of $E$ in $V$ is equal to $V$, then $E$ is absorbing in $V$. More precisely, if $| \cdot |$ is the trivial absolute value function on $E$, then $E$ is a linear subspace of $V$, and hence $E = V$. Otherwise, suppose that $| \cdot |$ is not the trivial absolute value function on $k$, and let $v \in V$ be given. Thus $v$ can be expressed as a linear combination of elements of $E$, by hypothesis. This implies that $t v \in E$ when $t \in k$ and $|t|$ is sufficiently small, because $E$ is a balanced subgroup of $V$.

Let us continue to suppose that $| \cdot |$ is not the trivial absolute value function on $k$. Let $A$ be a balanced subgroup of $V$ that is also absorbing in $V$, and let $N_A$ be as in (1.14.12). We would like to check that $N_A$ is a semi-ultranorm on $V$ with respect to $| \cdot |$ on $k$. We have already seen that $N_A$ satisfies the homogeneity condition (1.14.13), and so it is enough to show that $N_A$ satisfies the semi-ultranorm version of the triangle inequality. Let $v, w \in V$ be given, and let $r$ be a positive real number such that

\[ N_A(v), N_A(w) < r. \]  

This implies that there are nonzero elements $t_1(v), t_1(w)$ of $k$ such that

\[ |t_1(v)|, |t_1(w)| < r \]

and

\[ v \in t_1(v)A, \ w \in t_1(w)A, \]
by the definition (1.14.12) of $N_A$. Let us take $t_1$ to be equal to $t_1(v)$ or $t_1(w)$, in such a way that
\begin{equation}
|t_1| = \max(|t_1(v)|, |t_1(w)|).
\end{equation}
(1.15.5)

Thus $t_1(v) A, t_1(w) A \subseteq t_1 A$, because $A$ is balanced. This implies that $v$ and $w$ are both elements of $t_1 A$, by (1.15.4). It follows that
\begin{equation}
v + w \in t_1 A,
\end{equation}
(1.15.6)
because $A$ is a subgroup of $V$ with respect to addition. This means that
\begin{equation}
N_A(v + w) \leq |t_1|,
\end{equation}
(1.15.7)
by the definition (1.14.12) of $N_A$. Hence $N_A(v+w) < r$, by (1.15.3) and (1.15.5). This shows that
\begin{equation}
N_A(v + w) \leq \max(N_A(v), N_A(w)),
\end{equation}
(1.15.8)
since $r$ is any positive real number that satisfies (1.15.2).
Chapter 2

Some basic notions related to Lie algebras

2.1 Modules and homomorphisms

Let $k$ be a commutative ring with a multiplicative identity element $1 = 1_k$, and let $A$ be a commutative group, for which the group operations are expressed additively. Suppose that scalar multiplication on $A$ by elements of $k$ is defined, so that $ta$ is defined as an element of $A$ for every $t \in k$ and $a \in A$. If scalar multiplication satisfies the usual compatibility conditions with the group operations on $A$ and the ring operations on $k$, then $A$ is said to be a module over $k$. If $k$ is a field, then a module over $k$ is the same as a vector space over $k$. Any abelian group may be considered as a module over $\mathbb{Z}$, where $na$ is the sum of $n$ $a$’s in $A$ for each $a \in A$ and $n \in \mathbb{Z}^+$. Let $k$ be a commutative ring with a multiplicative identity element again, and let $A$ be a module over $k$. A submodule of $A$ is a subgroup $A_0$ of $A$ with respect to addition that is invariant under scalar multiplication by $k$. If $k$ is a field, then this is the same as a linear subspace.

Let $k$ be a field with an ultrametric absolute value function $| \cdot |$, and let $k_1$ be the closed unit ball in $k$ with respect to $| \cdot |$. Thus $k_1$ is a subring of $k$ that contains the multiplicative identity element in particular, as in Section 1.7. Let $V$ be a vector space over $k$, which may be considered as a module over $k_1$ as well. In this situation, a submodule of $V$ as a module over $k_1$ is the same as a balanced subgroup of $V$, as in Section 1.15.

Let $k$ be a commutative ring with a multiplicative identity element, and let $B$ be a module over $k$. If $X$ is a nonempty set, then the space of all functions on $X$ with values in $B$ is a module over $k$, with respect to pointwise addition and scalar multiplication of the functions on $X$.

Let $A$ be another module over $k$. A mapping $\phi$ from $A$ into $B$ is said to be a module homomorphism if $\phi$ is a group homomorphism with respect to addition that is also compatible with scalar multiplication by elements of $k$. One may
2.2. ALGEBRAS

say that $\phi$ is linear over $k$ or $k$-linear in this case as well. If $k$ is a field, then this is the same as a linear mapping between vector spaces.

The collection of all module homomorphisms from $A$ into $B$ may be denoted $\text{Hom}(A, B)$, or $\text{Hom}_k(A, B)$, to indicate the role of $k$. This is a module over $k$ too, with respect to pointwise addition and scalar multiplication of mappings from $A$ into $B$. More precisely, $\text{Hom}(A, B)$ may be considered as a submodule of the module of all $B$-valued functions on $A$.

Let $C$ be another module over $k$. If $\phi$ is a module homomorphism from $A$ into $B$, and $\psi$ is a module homomorphism from $B$ into $C$, then their composition $\psi \circ \phi$ defines a module homomorphism from $A$ into $C$.

If $\phi$ is a one-to-one module homomorphism from $A$ onto $B$, then the inverse mapping $\phi^{-1}$ is a module homomorphism from $B$ onto $A$. In this case, $\phi$ is said to be a module isomorphism from $A$ onto $B$. If $\phi$ is a module isomorphism from $A$ onto $B$, and $\psi$ is a module isomorphism from $B$ onto $C$, then $\psi \circ \phi$ is a module isomorphism from $A$ onto $C$.

A mapping $\beta$ from $A \times B$ into $C$ is said to be bilinear over $k$ if $\beta(a, b)$ is linear over $k$ in $a$ as a mapping from $A$ into $C$ for every $b \in B$, and $\beta(a, b)$ is linear over $k$ in $b$ as a mapping from $B$ into $C$ for every $a \in A$. If $k$ is a field, then this is the same as the usual notion of bilinearity for a mapping from a product of vector spaces over $k$ into another vector space over $k$.

In particular, we can take $A = B$, so that $\beta$ is a bilinear mapping from $A \times A$ into $C$. If

$$\beta(b, a) = \beta(a, b)$$

(2.1.1)

for every $a, b \in A$, then $\beta$ is said to be symmetric on $A \times A$. Similarly, if

$$\beta(b, a) = -\beta(a, b)$$

(2.1.2)

for every $a, b \in A$, then $\beta$ is said to be antisymmetric on $A \times A$. However, it is sometimes better to ask that

$$\beta(a, a) = 0$$

(2.1.3)

for every $a \in A$. Of course,

$$\beta(a + b, a + b) = \beta(a, a) + \beta(a, b) + \beta(b, a) + \beta(b, b)$$

(2.1.4)

for every $a, b \in A$, because of bilinearity. It is easy to see that (2.1.3) implies (2.1.2), using (2.1.4). If $1 + 1$ has a multiplicative inverse in $k$, then (2.1.2) implies (2.1.3). If $1 + 1 = 0$ in $k$, then (2.1.1) and (2.1.2) are the same.

2.2 Algebras

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be a module over $k$. If $A$ is equipped with a mapping from $A \times A$ into $A$ that is bilinear over $k$, then $A$ is said to be an algebra over $k$. In this case, we may also say that $A$ is an algebra over $k$ in the strict sense. The bilinear mapping may be expressed as

$$\beta(a, b) \mapsto ab,$$  

(2.2.1)
and we may refer to \(a b\) as the product of \(a, b \in A\). If this bilinear mapping is symmetric, so that
\[
(2.2.2) \quad ab = ba
\]
for every \(a, b \in A\), then we may say that \(A\) is commutative.

Let \(A\) be an algebra over \(k\). If the associative law
\[
(2.2.3) \quad (ab)c = a(bc)
\]
holds for every \(a, b, c \in A\), then \(A\) is said to be an associative algebra over \(k\). This is sometimes included in the definition of an algebra, and we do consider associativity to be part of the definition of a ring here.

Let \(A\) be an algebra over \(k\) in the strict sense. An element \(e\) of \(A\) is said to be the multiplicative identity element in \(A\) if
\[
(2.2.4) \quad ea = ae = a
\]
for every \(a \in A\). It is easy to see that this is unique when it exists.

Let \(A\) be any module over \(k\), so that the space \(\text{Hom}_k(A, A)\) of module homomorphisms from \(A\) into itself is a module over \(k\) too. One can check that \(\text{Hom}_k(A, A)\) is an associative algebra over \(k\), with composition of mappings as multiplication. The identity mapping on \(A\) is the multiplicative identity element in \(\text{Hom}_k(A, A)\).

Let \(A\) be an algebra over \(k\) in the strict sense, and let \(A_0\) be a submodule of \(A\), as a module over \(k\). If, for every \(a, b \in A_0\), we have that \(ab \in A_0\), then \(A_0\) is said to be a subalgebra of \(A\). In this case, \(A_0\) is also an algebra over \(k\) in the strict sense, with respect to the restriction of multiplication on \(A\) to \(A_0\). If \(A\) is an associative algebra, then \(A_0\) is associative as well.

Let \(A\) and \(B\) be algebras over \(k\) in the strict sense, so that \(A\) and \(B\) are modules over \(k\) in particular. Also let \(\phi\) be a module homomorphism from \(A\) into \(B\). If
\[
(2.2.5) \quad \phi(a_1a_2) = \phi(a_1)\phi(a_2)
\]
for every \(a_1, a_2 \in A\), then one may say that \(\phi\) is an algebra homomorphism from \(A\) into \(B\). If \(A\) and \(B\) have multiplicative identity elements \(e_A\) and \(e_B\), respectively, then one may require that
\[
(2.2.6) \quad \phi(e_A) = e_B
\]
too. If \(\phi\) is a one-to-one algebra homomorphism from \(A\) onto \(B\), then the inverse mapping \(\phi^{-1}\) is an algebra homomorphism from \(B\) onto \(A\), and \(\phi\) is said to be an algebra isomorphism from \(A\) onto \(B\). In this case, if \(A\) has a multiplicative identity element \(e_A\), then \(\phi(e_A)\) is the multiplicative identity element in \(B\). An algebra isomorphism from \(A\) onto itself is called an algebra automorphism of \(A\).

Let \(A\) be an algebra over \(k\) in the strict sense again. If \(a \in A\), then put
\[
(2.2.7) \quad M_a(x) = ax
\]
for every $x \in A$. Note that $M_a$ defines a module homomorphism from $A$ into itself for every $a \in A$, because of bilinearity of multiplication on $A$. Thus

$$a \mapsto M_a$$

may be considered as a mapping from $A$ into the space $\text{Hom}_k(A, A)$ of module homomorphisms from $A$ into itself. More precisely, (2.2.8) is a module homomorphism from $A$ into $\text{Hom}_k(A, A)$ as modules over $k$, because multiplication on $A$ is bilinear.

Suppose for the moment that $A$ has a multiplicative identity element $e$. In this case, $M_e$ is the identity mapping on $A$. We also get that

$$M_a(e) = a e = a$$

for every $a \in A$, which implies that (2.2.8) is injective.

If $a, b, x \in A$, then

$$M_a(M_b(x)) = M_a(b x) = a(b x)$$

and

$$M_{ab}(x) = (ab) x.$$
for every $x, y, z \in A$. Alternatively, the Jacobi identity may be expressed as saying that

\[ [x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \quad (2.3.4) \]

for every $x, y, z \in A$. It is easy to see that (2.3.3) is equivalent to (2.3.4), if we have (2.3.2). If $[x, y]$ is a bilinear mapping from $A \times A$ into $A$ that satisfies (2.3.1) and either (2.3.3) or (2.3.4), then $A$ is said to be a Lie algebra over $k$ with respect to the Lie bracket $[x, y]$. In particular, $A$ may be considered as an algebra over $k$ in the strict sense, using $[x, y]$ as multiplication on $A$.

If

\[ [x, y] = 0 \quad (2.3.5) \]

for every $x, y \in A$, then $A$ is said to be commutative as a Lie algebra. Commutativity of $A$ as an algebra over $k$ in the strict sense with respect to $[x, y]$ means that $[x, y]$ is symmetric in $x$ and $y$, as in the previous section. If $1 + 1$ is invertible in $k$, then commutativity of $A$ as an algebra in the strict sense implies that $A$ is commutative as a Lie algebra, because of (2.3.2). However, if $1 + 1 = 0$ in $k$, then any Lie algebra over $k$ is commutative as an algebra over $k$ in the strict sense.

Let $A$ be any algebra over $k$ in the strict sense, where multiplication on $A$ is expressed as in (2.2.1). Put

\[ [x, y] = xy - yx \quad (2.3.6) \]

for every $x, y \in A$, which defines a mapping from $A \times A$ into $A$. This mapping is bilinear over $k$, because multiplication on $A$ is bilinear, by hypothesis. Of course, (2.3.6) satisfies (2.3.1) automatically. If $A$ is associative with respect to the given operation of multiplication, then one can verify that (2.3.6) satisfies either of the Jacobi identities (2.3.3) or (2.3.4), so that $A$ is a Lie algebra with respect to (2.3.6).

Let $(A, [\cdot, \cdot]_A)$ and $(B, [\cdot, \cdot]_B)$ be Lie algebras over $k$. Thus $A$ and $B$ are modules over $k$ in particular, and we let $\phi$ be a module homomorphism from $A$ into $B$. If

\[ \phi([x, y]_A) = [\phi(x), \phi(y)]_B \quad (2.3.7) \]

for every $x, y \in A$, then $\phi$ is said to be a Lie algebra homomorphism from $A$ into $B$. Remember that $A$ and $B$ may be considered as algebras over $k$ in the strict sense, using $[\cdot, \cdot]_A$ and $[\cdot, \cdot]_B$ as multiplication on $A$ and $B$, respectively. A Lie algebra homomorphism from $A$ into $B$ is the same as an algebra homomorphism from $A$ into $B$, as algebras over $k$ in the strict sense, as in the previous section.
Let $A$ and $B$ be algebras over $k$ in the strict sense, with multiplication expressed as in (2.2.1), and let $\phi$ be an algebra homomorphism from $A$ into $B$. Also let $[,]_A$ and $[,]_B$ be the corresponding commutators on $A$ and $B$, as in (2.3.6). If $x, y \in A$, then

\begin{equation}
\phi([x, y]_A) = \phi(xy - yx) = \phi(x)\phi(y) - \phi(y)\phi(x) = [\phi(x), \phi(y)]_B.
\end{equation}

This means that $\phi$ may be considered as an algebra homomorphism from $A$ into $B$, using $[,]_A$ and $[,]_B$ as the algebra operations on $A$ and $B$, respectively.

If $A$ and $B$ are associative algebras with respect to their given operations of multiplication, then $A$ and $B$ are Lie algebras with respect to $[,]_A$ and $[,]_B$, respectively, as before, and $\phi$ may be considered as a Lie algebra homomorphism from $A$ into $B$ with respect to these Lie brackets.

### 2.4 The adjoint representation

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [x, y]_A)$ be a Lie algebra over $k$. If $x \in A$, then let $\text{ad} x$ be the mapping from $A$ into itself defined by

\begin{equation}
(\text{ad} x)(y) = [x, y]_A
\end{equation}

for every $y \in A$. This is the same as the multiplication operator on $A$ corresponding to $x$, as in Section 2.2, using the Lie bracket on $A$ as multiplication. In particular, $\text{ad} x$ is a module homomorphism from $A$ into itself, as a module over $k$, because $[x, y]_A$ is linear in $y$ over $k$. Thus

\begin{equation}
x \mapsto \text{ad} x
\end{equation}

defines a mapping from $A$ into the space $\text{Hom}_k(A, A)$ of module homomorphisms from $A$ into itself. Remember that $\text{Hom}_k(A, A)$ is a module over $k$ with respect to pointwise addition and scalar multiplication of mappings from $A$ into itself. As before, (2.4.2) is a module homomorphism from $A$ into $\text{Hom}_k(A, A)$ as modules over $k$, because $[x, y]_A$ is linear in $x$ over $k$.

If $\phi, \psi \in \text{Hom}_k(A, A)$, then put

\begin{equation}
[\phi, \psi] = [\phi, \psi]_{\text{Hom}_k(A, A)} = \phi \circ \psi - \psi \circ \phi,
\end{equation}

which is an element of $\text{Hom}_k(A, A)$ too. Of course, $\text{Hom}_k(A, A)$ is a Lie algebra over $k$ with respect to (2.4.3), because $\text{Hom}_k(A, A)$ is an associative algebra over $k$ with respect to composition of mappings. It is well known that (2.4.2) is a Lie algebra homomorphism from $A$ into $\text{Hom}_k(A, A)$, with respect to (2.4.3) on $\text{Hom}_k(A, A)$. To see this, let $x, y, z \in A$ be given, and observe that

\begin{equation}
[[\text{ad} x, \text{ad} y](z) = (\text{ad} x)((\text{ad} y)(z)) - (\text{ad} y)((\text{ad} x)(z))
= (\text{ad} x)([y, z]_A) - (\text{ad} y)([x, z]_A)
= [x, [y, z]_A]_A - [y, [x, z]_A]_A
= [x, [y, z]_A]_A + [y, [z, x]_A]_A.
\end{equation}
This uses the fact that $[x,z]_A = -[z,x]_A$, as in (2.3.2), in the last step. We also have that
(2.4.5) \[ (\text{ad}[x,y]_A)(z) = [[x,y]_A,z]_A = -[z,[x,y]_A]_A, \]
using (2.3.2) in the second step. The Jacobi identity (2.3.4) says exactly that the right sides of (2.4.4) and (2.4.5) are equal to each other. Thus
(2.4.6) \[ (\text{ad}[x,y]_A)(z) = ([\text{ad} x, \text{ad} y])(z) \]
for every $z \in A$, so that
(2.4.7) \[ \text{ad}[x,y]_A = [\text{ad} x, \text{ad} y] \]
as mappings from $A$ into itself.

2.5 Derivations

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an algebra over $k$ in the strict sense, where multiplication of $a,b \in A$ is expressed as $ab$. Also let $\delta$ be a module homomorphism from $A$ into itself, as a module over $k$. If $\delta$ satisfies the product rule
(2.5.1) \[ \delta(ab) = \delta(a) b + a \delta(b) \]
for every $a,b \in A$, then $\delta$ is said to be a derivation on $A$. Let Der($A$) be the collection of derivations on $A$. Remember that the space Hom$_k(A,A)$ of all module homomorphisms from $A$ into itself is a module over $k$ with respect to pointwise addition and scalar multiplication of mappings on $A$. It is easy to see that Der($A$) is a submodule of Hom$_k(A,A)$, as a module over $k$. We also have that Hom$_k(A,A)$ is an associative algebra over $k$, with respect to composition of mappings. This implies that Hom$_k(A,A)$ is a Lie algebra over $k$ with respect to the corresponding commutator bracket. It is well known that Der($A$) is a Lie subalgebra of Hom$_k(A,A)$ with respect to the commutator bracket.

More precisely, let $\delta, \delta' \in$ Der($A$) and $a,b \in A$ be given. Thus the commutator
(2.5.2) \[ [\delta, \delta'] = \delta \circ \delta' - \delta' \circ \delta \]
of $\delta$ and $\delta'$ is defined as a module homomorphism from $A$ into itself. Observe that
(2.5.3) \[ ([\delta, \delta'])(ab) = \delta(\delta'(a)b) - \delta'(\delta(a)b) = \delta(\delta'(a)b + a\delta'(b)) - \delta'(\delta(a)b + a\delta(b)) = \delta(\delta'(a)b + \delta'(a)\delta(b) + \delta(a)\delta'(b) + a\delta(\delta'(b)) - \delta'(\delta(a)b + \delta(a)\delta'(b) - \delta(a)\delta(b) - a\delta'(\delta(b))) \]
The middle pair of terms in the last two lines cancel each other, so that
(2.5.4) \[ ([\delta, \delta'])(ab) = \delta(\delta'(a)b + a\delta'(b)) - \delta'(\delta(a)b + a\delta'(b)) = ([\delta, \delta'])(a)b + a([\delta, \delta'])(b), \]
as desired.

Let \( a \in A \) be given, and put

\[
\delta_a(x) = ax - xa
\]

for every \( x \in A \). Of course, the right side of (2.5.5) is the same as the commutator bracket corresponding to multiplication on \( A \). Note that \( \delta_a \) is a module homomorphism from \( A \) into itself, because of bilinearity over \( k \) of multiplication on \( A \). Similarly,

\[
a \mapsto \delta_a
\]

is a module homomorphism from \( A \) into \( \text{Hom}_k(A, A) \), as modules over \( k \). If \( A \) is an associative algebra over \( k \) and \( x, y \in A \), then

\[
\delta_a(xy) = \delta(ax) - \delta(xy)a = (ax)y - x(ay) = \delta_a(x)y + x\delta_a(y),
\]

so that \( \delta_a \in \text{Der}(A) \).

Suppose now that \((A, [\cdot, \cdot]_A)\) is a Lie algebra over \( k \). Let \( \delta \) be a module homomorphism from \( A \) into itself, as a module over \( k \). In this situation, the product rule says that

\[
\delta([a, b]_A) = [\delta(a), b]_A + [a, \delta(b)]_A
\]

for every \( a, b \in A \). Thus \( \delta \in \text{Der}(A) \) when this holds. Let \( x \in A \) be given, and let us verify that \( \text{ad} x = [x, .]_A \in \text{Der}(A) \). If \( y, z \in A \), then

\[
(\text{ad} x)([y, z]_A) = [x, [y, z]_A]_A = -[y, [z, x]_A]_A - [z, [x, y]_A]_A,
\]

using the Jacobi identity in the second step. It follows that

\[
(\text{ad} x)([y, z]_A) = [[x, y]_A, z]_A + [y, [x, z]_A]_A
\]

as desired.

Let \( A \) be an algebra over \( k \) in the strict sense again, where the product of \( a, b \in A \) is denoted \( ab \). If \( x, y \in A \), then let \([x, y] = xy - yx\) be the usual commutator bracket corresponding to multiplication in \( A \). Let \( \delta \in \text{Der}(A) \) be given, and observe that

\[
\delta([x, y]) = \delta(xy - yx) = \delta(x)y + x\delta(y) - \delta(y)x - y\delta(x) = [\delta(x), y] + [x, \delta(y)]
\]

for every \( x, y \in A \). Of course, we can also consider \( A \) as an algebra over \( k \) in the strict sense with respect to \([x, y]\). It follows from (2.5.11) that \( \delta \) is a derivation on \( A \) with respect to \([x, y]\) as well.
2.6 Involutions

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$, $B$ be algebras over $k$ in the strict sense, where multiplication of $x, y$ is expressed as $xy$. In particular, $A$ and $B$ are modules over $k$, and we let $\phi$ be a module homomorphism from $A$ into $B$. If

\[
\phi(a_1 a_2) = \phi(a_2) \phi(a_1)
\]

for every $a_1, a_2 \in A$, then one may say that $\phi$ is an opposite algebra homomorphism from $A$ into $B$. Of course, this is the same as an ordinary algebra homomorphism when $A$ or $B$ is commutative. If $A$ and $B$ have multiplicative identity elements $e_A$ and $e_B$, respectively, then one may also ask that

\[
\phi(e_A) = e_B.
\]

If $\phi$ is a one-to-one opposite algebra homomorphism from $A$ onto $B$, then $\phi^{-1}$ is an opposite algebra homomorphism from $B$ onto $A$, and $\phi$ is said to be an opposite algebra isomorphism from $A$ onto $B$. In this case, if $A$ has a multiplicative identity element $e_A$, then it is easy to see that $\phi(e_B)$ is the multiplicative identity element in $B$. An opposite algebra automorphism on $A$ is an opposite algebra isomorphism from $A$ onto itself.

Let $[a_1, a_2]_A = a_1 a_2 - a_2 a_1$ and $[b_1, b_2]_B = b_1 b_2 - b_2 b_1$ be the corresponding commutator brackets on $A$ and $B$. If $\phi$ is an opposite algebra homomorphism from $A$ into $B$ and $a_1, a_2 \in A$, then

\[
\phi([a_1, a_2]_A) = \phi(a_1 a_2 - a_2 a_1) = \phi(a_2) \phi(a_1) - \phi(a_1) \phi(a_2) = -[\phi(a_1), \phi(a_2)]_B.
\]

An opposite algebra homomorphism $x \mapsto x^*$ from $A$ into itself is said to be an (algebra) involution on $A$ if

\[
(x^*)^* = x
\]

for every $x \in A$. This implies that $x \mapsto x^*$ is a one-to-one mapping from $A$ onto itself, which is its own inverse mapping.

Let $x \mapsto x^*$ be an opposite algebra automorphism on $A$. An element $a$ of $A$ is said to be self-adjoint with respect to $x \mapsto x^*$ if

\[
a^* = a,
\]

and $a$ is said to be anti-self-adjoint with respect to $x \mapsto x^*$ if

\[
a^* = -a.
\]

The collections of self-adjoint and anti-self-adjoint elements of $A$ are submodules of $A$, as a module over $k$. If $1 + 1 = 0$ in $k$, then self-adjointness and anti-self-adjointness are the same. If $1 + 1$ is invertible in $k$ and $a \in A$ is both self-adjoint and anti-self-adjoint, then $a = 0$. 

Suppose that $x \mapsto x^*$ is an algebra involution on $A$. If $a$ is any element of $A$, then

(2.6.7) \hspace{1cm} a + a^*

is self-adjoint with respect to $x \mapsto x^*$, and

(2.6.8) \hspace{1cm} a - a^*

is anti-self-adjoint with respect to $x \mapsto x^*$. If $1 + 1 = 0$ in $k$, then (2.6.7) and (2.6.8) are the same. If $1 + 1$ is invertible in $k$, then every element of $A$ can be expressed as the sum of self-adjoint and anti-self-adjoint elements of $A$, using (2.6.7) and (2.6.8). This expression is unique in this case, because $0$ is the only element of $A$ that is both self-adjoint and anti-self-adjoint.

Let $a, b \in A$ be given, and let $[a, b] = ab - ba$ be their usual commutator in $A$. If $x \mapsto x^*$ is an opposite algebra automorphism on $A$, and $a, b$ are anti-self-adjoint with respect to $x \mapsto x^*$, then

(2.6.9) \hspace{1cm} ([a, b])^* = -[a^*, b^*] = -[-a, -b] = -[a, b],

using (2.6.3) in the first step. Thus $[a, b]$ is anti-self-adjoint as well.

Suppose now that $k$ is the field $\mathbb{C}$ of complex numbers, and that $A$ and $B$ are algebras in the strict sense over $\mathbb{C}$. A conjugate-linear mapping $\phi$ from $A$ into $B$ is said to be a conjugate-linear algebra homomorphism if it preserves products as in (2.2.5), and $\phi$ is said to be a conjugate-linear opposite algebra homomorphism if it satisfies (2.6.1). If $A$ and $B$ are considered as algebras over the real numbers, then $\phi$ may be considered as a real-linear algebra homomorphism or opposite algebra homomorphism from $A$ into $B$, as appropriate. If $\phi$ is a one-to-one conjugate-linear algebra homomorphism or opposite algebra homomorphism from $A$ onto $B$, then $\phi^{-1}$ is a conjugate-linear algebra homomorphism or opposite algebra homomorphism from $B$ onto $A$, as appropriate, and $\phi$ is said to be a conjugate-linear algebra isomorphism or opposite algebra isomorphism from $A$ onto $B$, as appropriate. In particular, if $A = B$, then $\phi$ is said to be a conjugate-linear algebra automorphism or opposite algebra automorphism on $A$, as appropriate.

A conjugate-linear algebra homomorphism $x \mapsto x^*$ from $A$ into itself is said to be a conjugate-linear (algebra) involution on $A$ if it satisfies (2.6.4) for every $x \in A$. In this case, $x \mapsto x^*$ is a conjugate-linear opposite algebra automorphism on $A$, as before. Suppose that $x \mapsto x^*$ is a conjugate-linear opposite algebra automorphism on $A$, which may be considered as a real-linear opposite algebra automorphism of $A$ as an algebra over $\mathbb{R}$. In particular, $A$ may be considered as a vector space over $\mathbb{R}$, and the collections of self-adjoint and anti-self-adjoint elements of $A$ are real-linear subspaces of $A$, which is to say that they are linear subspaces of $A$ as a vector space over $\mathbb{R}$. In this situation, the anti-self-adjoint elements of $A$ are exactly those that can be expressed as $i$ times a self-adjoint element of $A$. 
2.7 More on multiplication operators

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an algebra over $k$ in the strict sense again, where multiplication of $a, b \in A$ is expressed as $a \cdot b$. If $a \in A$, then

$$M_a(x) = a \cdot x \quad (2.7.1)$$

defines a module homomorphism from $A$ into itself, as in Section 2.2. This is the operator of left multiplication by $a$ on $A$. Similarly,

$$\overline{M}_a(x) = x \cdot a \quad (2.7.2)$$

defines a module homomorphism from $A$ into itself, which is the operator of right multiplication by $a$ on $A$. Of course, if $A$ is commutative, then (2.7.1) and (2.7.2) are the same.

As before,

$$a \mapsto \overline{M}_a \quad (2.7.3)$$

defines a mapping from $A$ into the space $\text{Hom}_k(A, A)$ of all homomorphisms from $A$ into itself, as a module over $k$. Bilinearity of multiplication on $A$ implies that (2.7.3) is a module homomorphism from $A$ into $\text{Hom}_k(A, A)$, as modules over $k$. If $A$ has a multiplicative identity element $e$, then $e \mapsto \overline{M}_e$ is the identity mapping on $A$. We also have that

$$\overline{M}_a(e) = e \cdot a = a \quad (2.7.4)$$

for every $a \in A$ in this case, so that (2.7.3) is injective.

Observe that

$$\overline{M}_a(\overline{M}_b(x)) = \overline{M}_a(x \cdot b) = (x \cdot b) \cdot a \quad (2.7.5)$$

and

$$\overline{M}_b(x) = x \cdot (b \cdot a) \quad (2.7.6)$$

for every $a, b, x \in A$. If $A$ is an associative algebra, then we get that

$$\overline{M}_a \circ \overline{M}_b = \overline{M}_b \circ \overline{M}_a \quad (2.7.7)$$

for every $a, b \in A$, as mappings from $A$ into itself. This implies that (2.7.3) is an opposite algebra homomorphism from $A$ into $\text{Hom}_k(A, A)$, using composition of mappings as multiplication on $\text{Hom}_k(A, A)$, as usual.

Let $[x, y] = x \cdot y - y \cdot x$ be the usual commutator of $x, y \in A$. If $A$ is an associative algebra, then it follows that

$$\overline{M}_{[a, b]} = \overline{M}_{a \cdot b - b \cdot a} = \overline{M}_{a \cdot b} - \overline{M}_{b \cdot a} = \overline{M}_b \circ \overline{M}_a - \overline{M}_a \circ \overline{M}_b \quad (2.7.8)$$

for every $a, b \in A$. Similarly,

$$M_{[a, b]} = M_{a \cdot b - b \cdot a} = M_{a \cdot b} - M_{b \cdot a} = M_a \circ M_b - M_b \circ M_a \quad (2.7.9)$$
for every $a, b \in A$ in this situation, using (2.2.12) in the third step. If $a, b, x \in A$, then

\begin{equation}
M_a(\widetilde{M}_b(x)) = M_a(xb) = a \cdot (xb)
\end{equation}

and

\begin{equation}
\widetilde{M}_b(M_a(x)) = \widetilde{M}_b(ax) = (ax)b.
\end{equation}

These are the same when $A$ is an associative algebra, in which case

\begin{equation}
M_a \circ \widetilde{M}_b = \widetilde{M}_b \circ M_a
\end{equation}

for every $a, b \in A$.

If $a \in A$, then let $ad_a$ be the mapping from $A$ into itself defined by

\begin{equation}
(ad_a)(x) = [a, x] = a \cdot x - x \cdot a = M_a(x) - \widetilde{M}_a(x)
\end{equation}

for every $x \in A$. Equivalently,

\begin{equation}
ad_a = M_a - \widetilde{M}_a,
\end{equation}

which is a module homomorphism from $A$ into itself, as a module over $k$. We also have that

\begin{equation}
a \mapsto ad_a
\end{equation}

is a module homomorphism from $A$ into $\text{Hom}_k(A, A)$, as modules over $k$. If $A$ is an associative algebra, then

\begin{equation}
ad[a, b] = [ad_a, ad_b] = (ad_a) \circ (ad_b) - (ad_b) \circ (ad_a)
\end{equation}

for every $a, b \in A$, as mappings from $A$ into itself. More precisely, if $A$ is an associative algebra over $k$, then $A$ is also a Lie algebra over $k$ with respect to the commutator bracket $[x, y]$. Thus (2.7.16) follows from the analogous statement for Lie algebras. Alternatively, one can use (2.7.14) to reduce to the properties (2.7.8), (2.7.9), and (2.7.12) of the multiplication operators $M_a$ and $\widetilde{M}_a$.

2.8 Matrices

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be a module over $k$. If $n \in \mathbb{Z}_+$, then we let $M_n(A)$ be the space of $n \times n$ matrices with entries in $A$. An element of $M_n(A)$ may be given as $a = (a_{j,l})$, where $a_{j,l} \in A$ for every $j, l = 1, \ldots, n$. It is easy to see that $M_n(A)$ is also a module over $k$, with respect to entrywise addition and scalar multiplication.

Suppose that $A$ is an algebra over $k$ in the strict sense, where multiplication of $x, y \in A$ is expressed as $xy$. If $a, b \in M_n(A)$, then their product $c = ab$ is defined as usual by

\begin{equation}
c_{j,r} = \sum_{l=1}^{n} a_{j,l} b_{l,r}
\end{equation}
for every $j, r = 1, \ldots, n$. It is easy to see that this is bilinear in $a$ and $b$ over $k$, so that $M_n(A)$ is an algebra in the strict sense over $k$ with respect to matrix multiplication. If $A$ is an associative algebra over $k$, then $M_n(A)$ is associative with respect to matrix multiplication.

Suppose for the moment that $A$ has a multiplicative identity element $e$. The identity matrix $I = I_n$ in $M_n(A)$ is the $n \times n$ matrix whose diagonal entries are equal to $e$, and whose other entries are equal to 0. This is the multiplicative identity element in $M_n(A)$.

If $a = (a_{j, l}) \in M_n(A)$, then the transpose $a^t = (a^t_{j, l}) \in M_n(A)$ is defined as usual by
\begin{equation}
(a^t)_{j, l} = a_{l, j}
\end{equation}
for every $j, l = 1, \ldots, n$. Note that the mapping $a \mapsto a^t$ from a matrix to its transpose defines a module homomorphism from $M_n(A)$ into itself, which is to say that it is linear over $k$. Let $x \mapsto x^*$ be an opposite algebra automorphism on $A$, as in Section 2.6. If $a = (a_{j, l}) \in M_n(A)$, then let $a^* = ((a^*)_{j, l}) \in M_n(A)$ be defined by
\begin{equation}
(a^*)_{j, l} = (a_{l, j})^*
\end{equation}
for every $j, l = 1, \ldots, n$, which is to say that we apply $x \mapsto x^*$ to the entries of the transpose $a^t$ of $a$. One can check that this defines an opposite algebra automorphism on $M_n(A)$, and an involution on $M_n(A)$ when $x \mapsto x^*$ is an involution on $A$.

If $A$ is a commutative algebra, then the identity mapping on $A$ defines an algebra involution on $A$. In this case, $a^*$ reduces to the transpose $a^t$ of $a$.

Suppose now that $k$ is the field $\mathbb{C}$ of complex numbers. Let $x \mapsto x^*$ be a conjugate-linear opposite algebra automorphism on $A$, as in Section 2.6. In this situation, $a \mapsto a^*$ is a conjugate-linear opposite algebra automorphism on $M_n(A)$, and a conjugate-linear involution on $M_n(A)$ when $x \mapsto x^*$ is a conjugate-linear involution on $A$.

In particular, we can take $A = \mathbb{C}$, as a commutative algebra over itself. Of course, complex-conjugation may be considered as a conjugate-linear involution on $\mathbb{C}$. If $a = (a_{j, l}) \in M_n(\mathbb{C})$, then let $a^* = ((a^*)_{j, l}) \in M_n(\mathbb{C})$ be defined by
\begin{equation}
(a^*)_{j, l} = \overline{a_{l, j}}
\end{equation}
for every $j, l = 1, \ldots, n$, which is the same as the complex-conjugate of the entries of the transpose $a^t$ of $a$. This defines a conjugate-linear involution on $M_n(\mathbb{C})$, as before.

\section{Traces of matrices}

Let $k$ be a commutative ring with a multiplicative identity element, and let $n$ be a positive integer. If $A$ is a module over $k$, then $M_n(A)$ is a module over $k$ as well, with respect to entrywise addition and scalar multiplication, as in the
2.9. TRACES OF MATRICES

previous section. If \( a = (a_{j,l}) \in M_n(A) \), then the trace of \( a \) is defined as an element of \( A \) as usual by

\[
\text{tr} a = \sum_{j=1}^{n} a_{j,j}.
\]

(2.9.1)

It is easy to see that this defines a homomorphism from \( M_n(A) \) into \( A \), as modules over \( k \). We also have that

\[
\text{tr} a^t = \text{tr} a
\]

(2.9.2)

for every \( a \in M_n(A) \), where the transpose \( a^t \) of \( a \) is defined as in the previous section.

Suppose that \( A \) is an algebra over \( k \) in the strict sense, where multiplication of \( x, y \in A \) is expressed as \( xy \). If \( a, b \in M_n(A) \), then the products \( ab \) and \( ba \) are defined as elements of \( M_n(A) \) as in the previous section. Observe that

\[
\text{tr}(ab) = \sum_{j=1}^{n} a_{j,l} b_{l,j}
\]

(2.9.3)

and

\[
\text{tr}(ba) = \sum_{j=1}^{n} b_{l,j} a_{j,l}.
\]

(2.9.4)

If \( A \) is a commutative algebra over \( k \), then we get that

\[
\text{tr}(ab) = \text{tr}(ba)
\]

(2.9.5)

for every \( a, b \in M_n(A) \). Equivalently, this means that

\[
\text{tr}(ab - ba) = 0
\]

(2.9.6)

for every \( a, b \in M_n(A) \).

The \( n \)th general linear algebra \( gl_n(A) \) with entries in \( A \) is defined as an algebra over \( k \) in the following way. As a module over \( k \), \( gl_n(A) \) is the same as \( M_n(A) \). We use the commutator \([a, b] = ab - ba\) as the bilinear operation on \( gl_n(A) \), where the products \( ab \) and \( ba \) are as defined in the previous section. This makes \( gl_n(A) \) into an algebra over \( k \) in the strict sense. If \( A \) is an associative algebra over \( k \), then \( M_n(A) \) is an associative algebra over \( k \) too, so that \( gl_n(A) \) is a Lie algebra over \( k \).

Put

\[
sl_n(A) = \{ a \in gl_n(A) : \text{tr} a = 0 \},
\]

which defines a submodule of \( gl_n(A) \) as a module over \( k \), or equivalently a submodule of \( M_n(A) \). If \( A \) is a commutative algebra over \( k \), then

\[
[a, b] \in sl_n(A)
\]

(2.9.7)

for every \( a, b \in gl_n(A) \), as in (2.9.6). In particular, this means that \( sl_n(A) \) is a subalgebra of \( gl_n(A) \) with respect to the commutator \( [a, b] \) when \( A \) is commutative. In this case, \( sl_n(A) \) is called the \( n \)th special linear algebra with entries in \( A \). If \( A \) is a commutative associative algebra over \( k \), then \( sl_n(A) \) is a Lie algebra over \( k \) with respect to \([a, b]\).
CHAPTER 2. SOME BASIC NOTIONS RELATED TO LIE ALGEBRAS

2.10 Vector spaces and linear mappings

Let $k$ be a field. If $V$ and $W$ are vector spaces over $k$, then the space $\mathcal{L}(V,W)$ of linear mappings from $V$ into $W$ is a vector space over $k$ with respect to pointwise addition and scalar multiplication. This is the same as the space $\text{Hom}_k(V,W)$ of module homomorphisms from $V$ into $W$, where $V$ and $W$ are considered as modules over $k$. Similarly, if $V$ is a vector space over $k$, then the space $\mathcal{L}(V) = \mathcal{L}(V,V)$ of linear mappings from $V$ into itself is an associative algebra over $k$ with respect composition of mappings.

The general linear algebra $gl(V)$ associated to a vector space $V$ over $k$ is defined as a Lie algebra over $k$ in the following way. As a vector space over $k$, $gl(V)$ is the same as $\mathcal{L}(V)$. If $T_1$ and $T_2$ are linear mappings from $V$ into itself, then their commutator

$$[T_1, T_2] = T_1 \circ T_2 - T_2 \circ T_1 \quad (2.10.1)$$

defines a linear mapping from $V$ into itself as well. This defines a bilinear operation on $gl(V)$, which we use to define the Lie bracket on $gl(V)$. This satisfies the requirements of a Lie algebra, because $\mathcal{L}(V)$ is an associative algebra over $k$ with respect to composition of linear mappings.

Suppose that $V$ is a finite-dimensional vector space over $k$, with dimension $n \geq 1$. Let $v_1, \ldots, v_n$ be a basis for $V$, as a vector space over $k$. Thus every $v \in V$ can be expressed in a unique way as

$$v = \sum_{i=1}^{n} t_i v_i \quad (2.10.2)$$

where $t_1, \ldots, t_n \in k$. Let $a = (a_{j,l})$ be an $n \times n$ matrix with entries in $k$. If $v \in V$ is as in (2.10.2), then put

$$T_a(v) = \sum_{j=1}^{n} \left( \sum_{l=1}^{n} a_{j,l} t_l \right) v_j \quad (2.10.3)$$

which defines an element of $V$. Of course, $T_a$ is a linear mapping from $V$ into itself, and

$$a \mapsto T_a \quad (2.10.4)$$

is a linear mapping from $M_n(k)$ into $\mathcal{L}(V)$. More precisely, (2.10.4) is a one-to-one mapping from $M_n(k)$ onto $\mathcal{L}(V)$, and

$$T_a \circ T_b = T_{ab} \quad (2.10.5)$$

for every $a, b \in M_n(k)$. This means that (2.10.4) is an algebra isomorphism from $M_n(k)$ onto $\mathcal{L}(V)$, with respect to matrix multiplication on $M_n(k)$, and composition of linear mappings on $V$. It follows that (2.10.4) is also a Lie algebra isomorphism from $gl_n(k)$ onto $gl(V)$, with respect to their corresponding commutator brackets.
2.11. IDEALS AND QUOTIENTS

If \( a = (a_{j,l}) \in M_n(k) \), then the trace of \( T_a \) is defined as an element of \( k \) by

\[
(2.10.6) \quad \text{tr} T_a = \text{tr} a = \sum_{j=1}^{n} a_{j,j},
\]

where \( \text{tr} a \) refers to the trace of \( a \) as a matrix, as in the previous section. This defines the trace \( \text{tr} T \) of every linear mapping \( T \) from \( V \) into itself, by the remarks in the preceding paragraph. Note that the trace is a linear mapping from \( \mathcal{L}(V) \) into \( k \). If \( T_1, T_2 \in \mathcal{L}(V) \), then

\[
(2.10.7) \quad \text{tr}(T_1 \circ T_2) = \text{tr}(T_2 \circ T_1),
\]

by (2.9.5) and (2.10.5). It is well known that the trace of \( T \in \mathcal{L}(V) \) does not depend on the choice of basis \( v_1, \ldots, v_n \) for \( V \), because of (2.9.5) or (2.10.7).

Put

\[
(2.10.8) \quad \text{sl}(V) = \{ T \in \mathfrak{gl}(V) : \text{tr} T = 0 \},
\]

which is a linear subspace of \( \mathfrak{gl}(V) \), or equivalently of \( \mathcal{L}(V) \). If \( T_1, T_2 \in \mathfrak{gl}(V) \), then

\[
(2.10.9) \quad \text{tr}[T_1, T_2] = 0,
\]

by (2.10.7), and hence

\[
(2.10.10) \quad [T_1, T_2] \in \text{sl}(V).
\]

by (2.10.9). In particular, \( \text{sl}(V) \) is a subalgebra of \( \mathfrak{gl}(V) \) as a Lie algebra with respect to the commutator bracket. This is the special linear algebra associated to \( V \). The mapping (2.10.4) defines a Lie algebra isomorphism from \( \text{sl}_n(k) \) onto \( \text{sl}(V) \).

2.11 Ideals and quotients

Let \( k \) be a commutative ring with a multiplicative identity element. If \( A \) and \( B \) are modules over \( k \) and \( \phi \) is a module homomorphism from \( A \) into \( B \), then the kernel of \( \phi \) is the set of \( a \in A \) such that \( \phi(a) = 0 \), as usual. Of course, this is a submodule of \( A \).

If \( A \) is a module over \( k \), and \( A_0 \) is a submodule of \( A \) as a module over \( k \), then the quotient \( A/A_0 \) can be defined as a module over \( k \) in the usual way. More precisely, one can consider the quotient \( A/A_0 \) initially as a commutative group with respect to addition, and check that scalar multiplication on \( A/A_0 \) by elements of \( k \) can be defined in a natural way. The corresponding quotient mapping is a module homomorphism from \( A \) onto \( A/A_0 \), with kernel equal to \( A_0 \).

Let \( A \) be an algebra over \( k \) in the strict sense, where multiplication of \( a, b \in A \) is expressed as \( a b \). Also let \( A_0 \) be a submodule of \( A \), as a module over \( k \). If

\[
(2.11.1) \quad a b \in A_0
\]

for every \( a \in A \) and \( b \in A_0 \), then \( A_0 \) is said to be a left ideal in \( A \). Similarly, if (2.11.1) holds for every \( a \in A_0 \) and \( b \in A \), then \( A_0 \) is said to be a right ideal.
in \( A \). If \( A_0 \) is both a left and right ideal in \( A \), then \( A \) is said to be a two-sided ideal in \( A \). Of course, if \( A \) is a commutative algebra, then left and right ideals in \( A \) are the same. If \( B \) is another algebra over \( k \) in the strict sense, and \( \phi \) is an algebra homomorphism from \( A \) into \( B \), then the kernel of \( \phi \) is a two-sided ideal in \( A \).

If \( A_0 \) is a submodule of \( A \), as a module over \( k \), then the quotient \( A/A_0 \) can be defined as a module over \( k \) too, as before. Let \( q_0 \) be the corresponding quotient mapping from \( A \) onto \( A/A_0 \). Thus

\[
(2.11.2) \quad (a, b) \mapsto q_0(a b)
\]

is bilinear over \( k \) as a mapping from \( A \times A \) into \( A/A_0 \). If \( A_0 \) is a left ideal in \( A \), then

\[
(2.11.3) \quad q_0(a b) = 0
\]

for every \( a \in A \) and \( b \in A_0 \), and (2.11.2) leads to a bilinear mapping from \( A \times (A/A_0) \) into \( A/A_0 \). More precisely, if \( a, b \in A \), then \( q_0(a b) \) only depends on \( a \) and \( q_0(b) \) in this case. Similarly, if \( A_0 \) is a right ideal in \( A \), then (2.11.3) holds for every \( a \in A_0 \) and \( b \in A \), and (2.11.2) leads to a bilinear mapping from \( (A/A_0) \times A \) into \( A/A_0 \). If \( A_0 \) is a two-sided ideal in \( A \), then (2.11.3) holds when either \( a \) or \( b \) is in \( A_0 \), so that \( q_0(a b) \) only depends on \( q_0(a) \) and \( q_0(b) \). In this situation, (2.11.2) leads to a bilinear mapping from \( (A/A_0) \times (A/A_0) \) into \( A/A_0 \), which makes \( A/A_0 \) into an algebra over \( k \) in the strict sense, for which the quotient mapping \( q_0 \) is an algebra homomorphism.

Suppose for the moment that \( A \) is an associative algebra over \( k \). If \( A_0 \) is a two-sided ideal in \( A \), then \( A/A_0 \) is an associative algebra over \( k \) as well. If \( A_0 \) is a left ideal in \( A \), then elements of \( A \) act on \( A/A_0 \) by multiplication on the left, as in the preceding paragraph. Associativity of multiplication on \( A \) implies that the action on \( A/A_0 \) by products of elements of \( A \) corresponds to the composition of the actions on \( A/A_0 \) of the individual elements of \( A \). Similarly, if \( A_0 \) is a right ideal in \( A \), then elements of \( A \) act on \( A/A_0 \) by multiplication on the right, with the appropriate relationship between products of elements of \( A \) and their actions on \( A/A_0 \).

Suppose now that \( (A, [\cdot, \cdot]_A) \) is a Lie algebra over \( k \), and let \( A_0 \) be a submodule of \( A \) as a module over \( k \). If

\[
(2.11.4) \quad [a, b]_A \in A_0
\]

for every \( a \in A \) and \( b \in A_0 \), then \( A_0 \) is said to be an ideal in \( A \) as a Lie algebra. This is equivalent to saying that \( A_0 \) is a left, right, or two-sided ideal in \( A \), as an algebra over \( k \) in the strict sense. In this case, if \( q_0 \) is the usual quotient mapping from \( A \) onto \( A_0 \), then \( q_0([a, b]_A) \) depends only on \( q_0(a) \) and \( q_0(b) \), as before. It is easy to see that \( A/A_0 \) is also a Lie algebra over \( k \) with respect to the Lie bracket \([\cdot, \cdot]_{A/A_0}\) obtained from \([\cdot, \cdot]_A\) in this way.

Let \( A \) be an associative algebra over \( k \) again, where the product of \( a, b \in A \) is denoted \( a b \). Let \( A_0 \) be a two-sided ideal in \( A \). Remember that \( A \) may also be considered as a Lie algebra over \( k \) with respect to the commutator bracket \([a, b]_A = a b - b a\). Under these conditions, \( A_0 \) may be considered an ideal in \( A \) as a Lie algebra with respect to \([a, b]_A\) as well.
2.12 Bilinear forms

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$, $C$ be modules over $k$. Also let $\beta$ be a mapping from $A \times A$ into $C$ that is bilinear over $k$. Of course, $k$ may be considered as a module over itself as well, using multiplication on $k$ as scalar multiplication. If $C = k$, as a module over itself, then $\beta$ is said to be a bilinear form on $A$, as a module over $k$. As before, $\beta$ is said to be symmetric on $A \times A$ when

\[(2.12.1) \quad \beta(b, a) = \beta(a, b)\]

for every $a, b \in A$, and antisymmetric on $A \times A$ when

\[(2.12.2) \quad \beta(b, a) = -\beta(a, b)\]

for every $a, b \in A$. It is sometimes better to ask that

\[(2.12.3) \quad \beta(a, a) = 0\]

for every $a \in A$, instead of (2.12.2). Remember that (2.12.3) implies (2.12.2), because of bilinearity, as in Section 2.1. If $1 + 1 = 0$ in $k$, then (2.12.1) and (2.12.2) are the same. If $1 + 1$ is invertible in $k$, then (2.12.2) implies (2.12.3).

If $\beta$ is any bilinear mapping from $A \times A$ into $C$, then

\[(2.12.4) \quad \beta(a, b) + \beta(b, a)\]

is symmetric on $A \times A$, and

\[(2.12.5) \quad \beta(a, b) - \beta(b, a)\]

is antisymmetric on $A \times A$. If $1 + 1 = 0$ in $k$, then (2.12.4) is the same as (2.12.5), and is equal to 0 when $a = b$. If $1 + 1$ is invertible in $k$, then every bilinear mapping from $A \times A$ into $C$ can be expressed as the sum of a symmetric bilinear mapping and an antisymmetric bilinear mapping. In this case, a bilinear mapping from $A \times A$ into $C$ that is both symmetric and antisymmetric on $A \times A$ is identically 0 on $A \times A$, which implies that the previous expression as a sum is unique.

Let $\beta$ be a bilinear mapping from $A \times A$ into $C$, and let $\phi$ be a module homomorphism from $A$ into itself. Let us say that $\phi$ is symmetric with respect to $\beta$ on $A$ if

\[(2.12.6) \quad \beta(\phi(a), b) = \beta(a, \phi(b))\]

for every $a, b \in A$. Similarly, let us say that $\phi$ is antisymmetric with respect to $\beta$ on $A$ if

\[(2.12.7) \quad \beta(\phi(a), b) = -\beta(a, \phi(b))\]

for every $a, b \in A$. The collections of module homomorphisms from $A$ into itself that are symmetric or antisymmetric with respect to $\beta$ are submodules of Hom$_k(A, A)$. If $1 + 1 = 0$ in $k$, then (2.12.6) and (2.12.7) are the same, as usual.
Let $\phi$ and $\psi$ be module homomorphisms from $A$ into itself, and let $[\phi, \psi] = \phi \circ \psi - \psi \circ \phi$ be their commutator with respect to composition. If $\phi$ and $\psi$ are both symmetric on $A$ with respect to $\beta$, then

$$\beta(([\phi, \psi])(a), b) = \beta(\phi(\psi(a)), b) - \beta(\psi(\phi(a)), b)$$

$$= \beta(a, \psi(\phi(b))) - \beta(a, \phi(\psi(b))) = -\beta(a, ([\phi, \psi](b)))$$

for every $a, b \in A$. This also works when $\phi$ and $\psi$ are both antisymmetric with respect to $\beta$ and using antisymmetry twice in each term in the second step. In both cases, we get that $[\phi, \psi]$ is antisymmetric with respect to $\beta$ on $A$. In particular, the collection of module homomorphisms from $A$ into itself that are antisymmetric with respect to $\beta$ is a Lie subalgebra of $\text{Hom}_k(A, A)$, as a Lie algebra over $k$ with respect to the commutator bracket.

### 2.13 Dual spaces and mappings

Let $k$ be a field, and let $V$ be a vector space over $k$. Remember that a linear functional on $V$ is a linear mapping from $V$ into $k$, where $k$ is considered as a one-dimensional vector space over itself. The dual $V'$ of $V$ is the space of all linear functionals on $V$, which is a vector space over $k$ with respect to pointwise addition and scalar multiplication. If $V$ has finite dimension, then it is well known that the dimension of $V'$ is the same as the dimension of $V$. This can be seen by expressing linear functionals on $V$ in terms of a basis for $V$.

Let $W$ be another vector space over $k$, and let $T$ be a linear mapping from $V$ into $W$. If $\mu$ is a linear functional on $W$, then

$$T'(\mu) = \mu \circ T$$

is a linear functional on $V$. This defines a linear mapping $T'$ from $W'$ into $V'$, which is the dual mapping associated to $T$. We also have that

$$T \mapsto T'$$

is linear as a mapping from the space $\mathcal{L}(V, W)$ of linear mappings from $V$ into $W$ into the space $\mathcal{L}(W', V')$ of linear mappings from $W'$ into $V'$. The dual of the identity mapping $I_V$ on $V$, as a linear mapping from $V$ into itself, is the identity mapping $I_{V'}$ on $V'$.

Let $Z$ be a third vector space over $k$, let $T_1$ be a linear mapping from $V$ into $W$, and let $T_2$ be a linear mapping from $W$ into $Z$. Thus the composition $T_2 \circ T_1$ is a linear mapping from $V$ into $Z$, whose dual maps $Z'$ into $V'$. If $\nu \in Z'$, then

$$(T_2 \circ T_1)'(\nu) = \nu \circ (T_2 \circ T_1) = (\nu \circ T_2) \circ T_1 = T_1'(T_2'(\nu)).$$

This shows that

$$T_2 \circ T_1' = T_1' \circ T_2'$$

as mappings from $Z'$ into $V'$.
Let $V'' = (V')'$ be the dual of $V'$. If $v \in V$ and $\lambda \in V'$, then

\begin{equation}
L_v(\lambda) = \lambda(v)
\end{equation}

is an element of $k$. This defines $L_v$ as a linear functional on $V'$, and

\begin{equation}
v \mapsto L_v
\end{equation}

is a linear mapping from $V$ into $V''$. If $v \in V$ and $v \neq 0$, then one can find a $\lambda \in V'$ such that $\lambda(v) \neq 0$, using a basis for $V$. This implies that (2.13.6) is injective as a mapping from $V$ into $V''$. If $V$ has finite dimension, then $V'$ has the same dimension as $V$, and hence $V''$ has the same dimension as well. In this case, (2.13.6) also maps $V$ onto $V''$.

Let $\lambda_1, \ldots, \lambda_n$ be $n$ linear functionals on $V$ for some positive integer $n$, and put

\begin{equation}
\Lambda(v) = (\lambda_1(v), \ldots, \lambda_n(v))
\end{equation}

for each $v \in V$. This defines a linear mapping from $V$ into the space $k^n$ of $n$-tuples of elements of $k$, which is a vector space over $k$ with respect to coordinatewise addition and scalar multiplication, as usual. Of course, the kernel of $\Lambda$ is the same as the intersection of the kernels of $\lambda_1, \ldots, \lambda_n$. If $\Lambda$ is injective, then the dimension of $V$ is less than or equal to $n$, by standard results in linear algebra. In particular, if $\Lambda$ is injective and the dimension of $V$ is equal to $n$, then $\Lambda$ maps $V$ onto $k^n$.

### 2.14 Nondegenerate bilinear forms

Let $k$ be a field, and let $V$ be a finite-dimensional vector space over $k$. Also let $b(v, w)$ be a bilinear form on $V$. If $w \in V$, then

\begin{equation}
b_w(v) = b(v, w)
\end{equation}

defines a linear functional on $V$ as a function of $v$, and $w \mapsto b_w$ defines a linear mapping from $V$ into its dual space $V'$. The image

\begin{equation}
\{b_w : w \in V\}
\end{equation}

of this linear mapping is a linear subspace of $V'$. Note that (2.14.2) is equal to $V'$ exactly when $w \mapsto b_w$ is injective as a linear mapping from $V$ into $V'$, because $V$ and $V'$ have the same dimension.

If for every $v \in V$ with $v \neq 0$ there is a $w \in V$ such that $b(v, w) \neq 0$, then $b$ is said to be nondegenerate on $V$. This is the same as saying that the intersections of the kernels of the $b_w$’s, $w \in V$, is the trivial subspace $\{0\}$ of $V$. One can check that this happens exactly when (2.14.2) is equal to $V'$.

Suppose that $b$ is a nondegenerate bilinear form on $V$, and let $T$ be a linear mapping from $V$ into itself. If $w \in V$, then $b(T(v), w)$ defines a linear functional
on $V$, as a function of $v$. This implies that there is a unique element $T^*(w)$ of $V$ such that

$$b(T(v), w) = b(v, T^*(w)) \quad (2.14.3)$$

for every $v \in V$, because $b$ is nondegenerate on $V$. This defines a mapping $T^*$ from $V$ into itself, which is the adjoint of $T$ with respect to $b$. It is easy to see that $T^*$ is a linear mapping from $V$ into itself, because $T^*(w)$ is uniquely determined by (2.14.3).

Remember that the space $\mathcal{L}(V)$ of linear mappings from $V$ into itself is an algebra over $k$ with respect to composition of mappings. One can check that $T \mapsto T^*$ defines a linear mapping from $\mathcal{L}(V)$ into itself, because $T^*$ is uniquely determined by (2.14.3). Clearly $I^* = I$, where $I$ is the identity mapping on $V$.

If $T$ is any linear mapping from $V$ into itself, then $T^* = T$ exactly when $T$ is symmetric with respect to $b$, as in Section 2.12. Similarly, $T^* = -T$ exactly when $T$ is antisymmetric with respect to $b$.

If $T$ is a linear mapping from $V$ into itself and $T^* = 0$ on $V$, then $T = 0$ on $V$, because of (2.14.3) and the nondegeneracy of $b$ on $V$. This implies that $T \mapsto T^*$ is a one-to-one mapping from $\mathcal{L}(V)$ onto itself, because $\mathcal{L}(V)$ is a finite-dimensional vector space over $k$.

Let $T_1$ and $T_2$ be linear mappings from $V$ into itself. If $v, w \in V$, then

$$b((T_2 \circ T_1)(v), w) = b(T_2(T_1(v)), w) = b(T_1(v), T_2^*(w)) = b(v, T_1^*(T_2^*(w))) = b(v, (T_1^* \circ T_2^*)(w)). \quad (2.14.4)$$

This implies that

$$\quad (T_2 \circ T_1)^* = T_1^* \circ T_2^*, \quad (2.14.5)$$

so that $T \mapsto T^*$ is an opposite algebra automorphism on $\mathcal{L}(V)$, as in Section 2.6.

Let $T$ be a linear mapping from $V$ into itself again, so that $T^*$ and hence $(T^*)^*$ are defined as linear mappings from $V$ into itself, as before. If $b$ is symmetric on $V$, then

$$b(T(v), w) = b(T^*(w), v) = b(w, (T^*)^*(v)) = b((T^*)^*(v), w) \quad (2.14.6)$$

for every $v, w \in V$. Similarly, if $b$ is antisymmetric on $V$, then

$$b(T(v), w) = -b(T^*(w), v) = -b(w, (T^*)^*(v)) = b((T^*)^*(v), w) \quad (2.14.7)$$

for every $v, w \in V$. In both cases, we get that

$$\quad (T^*)^* = T. \quad (2.14.8)$$

It follows that $T \mapsto T^*$ defines an involution on $\mathcal{L}(V)$, as an algebra over $k$ with respect to composition, when $b$ is symmetric or antisymmetric on $V$.  

2.15 Sesquilinear forms

Let $V$ be a vector space over the field $\mathbb{C}$ of complex numbers. A complex-valued function $b$ on $V \times V$ is said to be sesquilinear if $b(v,w)$ is complex-linear in $v$ for each $w \in V$, and $b(v,w)$ is conjugate-linear in $w$ for every $v \in V$. In particular, if we consider $V$ and $\mathbb{C}$ as vector spaces over the real numbers, then it follows that $b$ is bilinear over $\mathbb{R}$. If we also have that

\[ b(v, w) = \overline{b(w, v)} \tag{2.15.1} \]

for every $v, w \in V$, then $b$ is said to be Hermitian-symmetric on $V$, or equivalently $b$ is a Hermitian form on $V$. The analogous Hermitian-antisymmetry condition

\[ b(v, w) = -\overline{b(w, v)} \tag{2.15.2} \]

is the same as saying that $ib(v, w)$ is a Hermitian form on $V$. If $b$ is any sesquilinear form on $V$, then $\overline{b(w, v)}$ is a sesquilinear form on $V$ too,

\[ b(v, w) + \overline{b(w, v)} \tag{2.15.3} \]

is a Hermitian form on $V$, and

\[ b(v, w) - \overline{b(w, v)} \tag{2.15.4} \]

is Hermitian-antisymmetric on $V$. This permits us to express $b(v, w)$ in a unique way as $b_1(v, w) + ib_2(v, w)$, where $b_1(v, w)$ and $b_2(v, w)$ are Hermitian forms on $V$. Note that $b(v, v) \in \mathbb{R}$ for every $v \in V$ when $b$ is a Hermitian form on $V$.

Let $b$ be a sesquilinear form on $V$, and let $T$ be a linear mapping from $V$ into itself. Let us say that $T$ is self-adjoint with respect to $b$ on $V$ if

\[ b(Tv, w) = b(v, Tw) \tag{2.15.5} \]

for every $v, w \in V$, and that $T$ is anti-self-adjoint with respect to $b$ on $V$ if

\[ b(Tv, w) = -b(v, Tw) \tag{2.15.6} \]

for every $v, w \in V$. One can check that $T$ is anti-self-adjoint with respect to $b$ on $V$ if and only if $iT$ is self-adjoint with respect to $b$ on $V$. If we consider $b$ as a real-bilinear mapping from $V \times V$ into $\mathbb{C}$, then these self-adjointness and anti-self-adjointness conditions correspond exactly to the symmetry and anti-symmetry conditions for module homomorphisms mentioned in Section 2.12.

The space of self-adjoint linear mappings from $V$ into itself with respect to $b$ on $V$ is a real-linear subspace of the space $\mathcal{L}(V)$ of linear mappings from $V$ into itself, which is to say that it is a linear subspace of $\mathcal{L}(V)$ when $\mathcal{L}(V)$ is considered as a vector space over $\mathbb{R}$. If $T_1, T_2$ are self-adjoint linear mappings from $V$ into itself with respect to $b$, then their commutator $[T_1, T_2] = T_1 \circ T_2 - T_2 \circ T_1$ with respect to composition is anti-self-adjoint with respect to $b$, as before. Similarly, if $T_1, T_2$ are anti-self-adjoint linear mappings from $V$ into itself with respect to $b$, then $[T_1, T_2]$ is anti-self-adjoint with respect to $b$ as well. It follows that the
space of anti-self-adjoint linear mappings from \( V \) into itself with respect to \( b \) is a real Lie subalgebra of \( \mathcal{L}(V) \), which is to say that it is a subalgebra of \( \mathcal{L}(V) \) as a Lie algebra over \( \mathbb{R} \) with respect to the commutator bracket.

Suppose from now on in this section that \( V \) has finite dimension as a complex vector space. If \( w \in V \), then

\[
(2.15.7) \quad b_w(v) = b(v, w)
\]

defines a linear functional on \( V \) as a function of \( v \), as before. In this situation, \( w \mapsto b_w \) is a conjugate-linear mapping from \( V \) into its dual space \( V' \). The image

\[
(2.15.8) \quad \{b_w : w \in V\}
\]
of this mapping is still a linear subspace of \( V' \), as a complex vector space. One can check that (2.15.8) is equal to \( V' \) exactly when \( w \mapsto b_w \) is injective as a mapping from \( V \) into \( V' \), because \( V \) and \( V' \) have the same dimension as complex vector spaces.

If for every \( v \in V \) with \( v \neq 0 \) there is a \( w \in V \) such that \( b(v, w) \neq 0 \), then \( b \) is said to be nondegenerate as a sesquilinear form on \( V \). This is the same as saying that the intersections of the kernels of the \( b_w \)'s, \( w \in V \), is trivial, as before. This happens exactly when (2.15.8) is equal to \( V' \), as in the previous section.

Let \( b \) be a nondegenerate sesquilinear form on \( V \), and let \( T \) be a linear mapping from \( V \) into itself. Also let \( w \in V \) be given, so that \( b(T(v), w) \) defines a linear functional on \( V \), as a function of \( v \). It follows that there is a unique element \( T^*(w) \) of \( V \) such that

\[
(2.15.9) \quad b(T(v), w) = b(v, T^*(w))
\]

for every \( v \in V \), because \( b \) is nondegenerate on \( V \). The resulting mapping \( T^* \) from \( V \) into itself is called the adjoint of \( T \) with respect to \( b \). One can check that \( T^* \) is a linear mapping from \( V \) into itself, using the sesquilinearity of \( b \).

However, \( T \mapsto T^* \) is conjugate-linear as a mapping from \( \mathcal{L}(V) \) into itself in this situation. We still have that \( I^* = I \), where \( I \) is the identity mapping on \( V \). A linear mapping \( T \) from \( V \) into itself is self-adjoint with respect to \( b \) if and only if \( T^* = T \). Similarly, \( T \) is anti-self-adjoint with respect to \( b \) if and only if \( T^* = -T \).

If \( T \) is a linear mapping from \( V \) into itself and \( T^* = 0 \) on \( V \), then \( T = 0 \) on \( V \), because of (2.15.9) and nondegeneracy of \( b \) on \( V \), as before. If \( T_1 \) and \( T_2 \) are linear mappings from \( V \) into itself, then one can verify that

\[
(2.15.10) \quad (T_2 \circ T_1)^* = T_1^* \circ T_2^* ,
\]
in the same way as before. It follows that \( T \mapsto T^* \) is a conjugate-linear opposite algebra automorphism on \( \mathcal{L}(V) \), because \( \mathcal{L}(V) \) is a finite-dimensional vector space over \( \mathbb{C} \).
Suppose now that $b$ is also Hermitian-symmetric on $V$. Let $T$ be a linear mapping from $V$ into itself, so that $T^*$ and $(T^*)^*$ are defined as linear mappings from $V$ into itself as well. If $v, w \in V$, then

\[(2.15.11) \quad b(T(v), w) = \overline{b(T^*(w), v)} = \overline{b(w, (T^*)^*(v))} = b((T^*)^*(v), w).\]

Thus

\[(2.15.12) \quad (T^*)^* = T,\]

so that $T \mapsto T^*$ is a conjugate-linear involution on $\mathcal{L}(V)$ in this case.
Chapter 3

Submultiplicativity and invertibility

3.1 Invertibility

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an associative algebra over $k$ with a multiplicative identity element $e$, where multiplication of $a, b \in A$ is expressed as $ab$. An element $a$ of $A$ is said to be invertible in $A$ if there is an element $b$ of $A$ such that

(3.1.1) \[ a b = b a = e. \]

It is easy to see that $b$ is unique when it exists, using associativity of multiplication on $A$. In this case, $b$ is called the multiplicative inverse of $a$ in $A$, and is denoted $a^{-1}$. Of course, $e$ is its own inverse in $A$. If $x$ and $y$ are invertible elements of $A$, then $xy$ is invertible in $A$ too, with

(3.1.2) \[ (xy)^{-1} = y^{-1} x^{-1}. \]

Thus the collection of invertible elements in $A$ is a group.

Let $x$ and $y$ be commuting elements of $A$, so that $xy = yx$. If $x$ is invertible in $A$, then $x^{-1}$ commutes with $y$ too. Suppose that $w$ and $z$ are commuting elements of $A$, and $wz$ is invertible in $A$. Note that $wz$ commutes with $w$ and $z$, so that $(wz)^{-1}$ commutes with $w$ and $z$ too. It follows that $w$ and $z$ are invertible in $A$, with $w^{-1} = (wz)^{-1} z$ and $z^{-1} = (wz)^{-1} w$.

Let $a \in A$ be given, and let $n$ be a nonnegative integer. Using a standard computation, we get that

(3.1.3) \[ (e - a) \sum_{j=0}^{n} a^j = \left( \sum_{j=0}^{n} a^j \right) (e - a) = e - a^{n+1}, \]

where $a^0$ is interpreted as being equal to $e$ when $j = 0$. In particular, if $a^{n+1} = 0,$

50
then it follows that $e - a$ is invertible in $A$, with

\[(e - a)^{-1} = \sum_{j=0}^{n} a^j.\]  

If $e - a^{n+1}$ is invertible in $A$, then (3.1.3) implies that $e - a$ is invertible in $A$ too, as in the previous paragraph.

Let $n$ be a positive integer, and consider the algebra $M_n(A)$ of $n \times n$ matrices with entries in $A$. The group of invertible elements of $M_n(A)$ is denoted $GL_n(A)$, and is called the $n$th general linear group with entries in $A$.

Suppose for the moment that $A$ is also commutative, so that the determinant of $a = (a_{ij}) \in M_n(A)$ can be defined as an element of $A$ in the usual way. If $a \in GL_n(A)$, then $\det a$ is an invertible element of $A$. Conversely, if $a \in M_n(A)$ and $\det a$ is an invertible element of $A$, then $a \in GL_n(A)$, by Cramer’s rule. The $n$th special linear group $SL_n(A)$ with entries in $A$ consists of the $a \in M_n(A)$ such that $\det a$ is the multiplicative identity element $e$ in $A$. This is a normal subgroup of $GL_n(A)$, because $SL_n(A)$ is the kernel of the determinant as a group homomorphism from $GL_n(A)$ into the multiplicative group of invertible elements in $A$.

Let $k$ be a field, and let $V$ be a vector space over $k$. Remember that the space $L(V)$ of linear mappings from $V$ into itself is an associative algebra with respect to composition of mappings, and with the identity mapping $I = I_V$ on $V$ as the multiplicative identity element in $L(V)$. The group $GL(V)$ of one-to-one linear mappings from $V$ onto itself with respect to composition of mappings is the same as the group of invertible elements in $L(V)$, and may be called the general linear group associated to $V$.

Suppose that $V$ has finite dimension $n \geq 1$, and let $v_1, \ldots, v_n$ be a basis for $V$. This leads to an algebra isomorphism from $M_n(k)$ onto $L(V)$, as in Section 2.10. The restriction of this mapping to $GL_n(k)$ defines a group isomorphism from $GL_n(k)$ onto $GL(V)$.

The determinant of a linear mapping $T$ from $V$ into itself can be defined as an element of $k$ as the determinant of the corresponding matrix in $M_n(k)$. It is well known that this does not depend on the choice of basis $v_1, \ldots, v_n$ of $V$. Note that $a \in M_n(k)$ is invertible exactly when $\det a \neq 0$, because $k$ is a field. Thus $T \in L(V)$ is invertible exactly when $\det T \neq 0$.

The special linear group $SL(V)$ associated to $V$ consists of the linear mappings $T$ from $V$ into itself such that $\det T = 1$ in $k$. In particular, these linear mappings are invertible on $V$, and $SL(V)$ is a normal subgroup of $GL(V)$, because it is the kernel of the determinant as a group homomorphism from $GL(V)$ into the multiplicative group of non-zero elements of $k$. The restriction of the algebra isomorphism from $M_n(k)$ onto $L(V)$ mentioned earlier to $SL_n(k)$ defines a group isomorphism from $SL_n(k)$ onto $SL(V)$. 

The determinant of $T$ is a well-defined element of $k$. If $T$ is invertible, then $\det T \neq 0$ in $k$. If $T$ is not invertible, then $\det T = 0$ in $k$. Since $SL(V)$ is the kernel of the determinant, it follows that $T \in SL(V)$ if and only if $\det T = 1$ in $k$.
3.2 Submultiplicative seminorms

Let $k$ be a field with an absolute value function $|\cdot|$, and let $A$ be an algebra over $k$ in the strict sense, where multiplication of $a,b \in A$ is expressed as $ab$. Also let $N_A$ be a seminorm on $A$, as a vector space over $k$, and with respect to $|\cdot|$ on $k$. As in Section 1.13, multiplication on $A$ is bounded as a bilinear mapping from $A \times A$ into $A$ with respect to $N_A$ on $A$ if there is a nonnegative real number $C$ such that

\[
N_A(ab) \leq C N_A(a) N_A(b)
\]

for every $a,b \in A$. If this holds with $C = 1$, then $N_A$ is said to be submultiplicative on $A$. Similarly, if

\[
N_A(ab) = N_A(a) N_A(b)
\]

for every $a,b \in A$, then $N_A$ is said to be multiplicative on $A$.

Suppose for the moment that $A$ has a multiplicative identity element $e$. If (3.2.1) holds for some $C \geq 0$, then we get that

\[
N_A(a) \leq C N_A(a) N_A(e)
\]

for every $a \in A$. If $N_A(a) > 0$ for some $a \in A$, then it follows that

\[
1 \leq C N_A(e).
\]

Let $V$ be a vector space over $k$, and let $N_V$ be a seminorm on $V$ with respect to $|\cdot|$ on $k$. Consider the space $\mathcal{B}(V) = \mathcal{B}(V,V)$ of bounded linear mappings from $V$ into itself, with respect to $N_V$ on $V$. This is a subalgebra of the algebra $\mathcal{L}(V)$ of all linear mappings from $V$ into itself, with composition of mappings as multiplication. Let $\|\cdot\|_{op} = \|\cdot\|_{op,VV}$ be the operator seminorm on $\mathcal{B}(V)$ corresponding to $N_V$ on $V$, as in Section 1.9. This is a submultiplicative seminorm on $\mathcal{B}(V)$ with respect to $|\cdot|$ on $k$, as before. It is easy to see that the identity mapping $I = I_V$ on $V$ is bounded with respect to $N_V$, with

\[
\|I\|_{op} = 1
\]

when $N_V(v) > 0$ for some $v \in V$, and $\|I\|_{op} = 0$ otherwise.

Let $A$ be an algebra over $k$ in the strict sense again, and let $N_A$ be a seminorm on $A$ with respect to $|\cdot|$ on $k$ that satisfies (3.2.1) for some $C \geq 0$. If $a \in A$, then

\[
M_a(x) = ax
\]

defines a linear mapping from $A$ into itself, as a vector space over $k$, as in Section 2.2. Using (3.2.1), we get that

\[
N_A(M_a(x)) \leq C N_A(a) N_A(x)
\]

for every $x \in A$, so that $M_a$ is bounded as a linear mapping from $A$ into itself with respect to $N_A$. More precisely, we have that

\[
\|M_a\|_{op} \leq C N_A(a)
\]
for every $a \in A$, where $\| \cdot \|_{op} = \| \cdot \|_{op, AA}$ is the operator seminorm on the space $\mathcal{B}(A)$ of bounded linear mappings from $A$ into itself with respect to $N_A$. If $A$ has a multiplicative identity element $e$, then

\begin{equation}
N_A(a) = N_A(M_a(e)) \leq \|M_a\|_{op} N_A(e)
\end{equation}

for every $a \in A$.

Similarly, if $a \in A$, then

\begin{equation}
\tilde{M}_a(x) = xa
\end{equation}

defines a linear mapping from $A$ into itself, as in Section 2.7. As before, we can use (3.2.1) to get that

\begin{equation}
N_A(M_a(x)) \leq C N_A(x) N_A(a)
\end{equation}

for every $x \in A$. This implies that $\tilde{M}_a$ is bounded as a linear mapping from $A$ into itself with respect to $N_A$, with

\begin{equation}
\|\tilde{M}_a\|_{op} \leq C N_A(a)
\end{equation}

for every $a \in A$. If $A$ has a multiplicative identity element $e$, then

\begin{equation}
N_A(a) = N_A(M_e(e)) \leq \|\tilde{M}_a\|_{op} N_A(e)
\end{equation}

for every $a \in A$.

Let $\alpha$ be a positive real number, and put

\begin{equation}
\tilde{N}_A(a) = \alpha N_A(a)
\end{equation}

for every $a \in A$. This defines a seminorm on $A$ as a vector space over $k$ too, with respect to $| \cdot |$ on $k$. Using (3.2.1), we get that

\begin{equation}
\tilde{N}_A(a b) \leq (C/\alpha) \tilde{N}_A(a) \tilde{N}_A(b)
\end{equation}

for every $a, b \in A$. In particular, this means that $\tilde{N}_A$ is submultiplicative on $A$ when $\alpha \geq C$.

Alternatively, $\|M_a\|_{op}$ defines a seminorm on $A$ as a vector space over $k$, with respect to $| \cdot |$ on $k$. Suppose that $A$ is an associative algebra over $k$, so that $a \mapsto M_a$ is an algebra homomorphism from $A$ into $\mathcal{B}(A)$, as in Section 2.2. This implies that $\|M_a\|_{op}$ is submultiplicative as a seminorm on $A$, because $\| \cdot \|_{op}$ is submultiplicative on $\mathcal{B}(A)$. If $A$ has a multiplicative identity element $e$, then $M_e$ is the identity operator on $A$, as before. Hence $\|M_e\|_{op}$ is equal to 1 when $N_A(a) > 0$ for some $a \in A$, and is equal to 0 otherwise.

### 3.3 Some matrix seminorms

Let $k$ be a field with an absolute value function $| \cdot |$, and let $A$ be an algebra over $k$ in the strict sense again, where multiplication of $a, b \in A$ is expressed as $a b$. 


Also let \( n \) be a positive integer, and let \( M_n(A) \) be the space of \( n \times n \) matrices with entries in \( A \), which is an algebra over \( k \) in the strict sense with respect to matrix multiplication, as in Section 2.8. Suppose that \( N_A \) is a seminorm on \( A \) with respect to \( | \cdot | \) on \( k \) that satisfies (3.2.1) for some \( C \geq 0 \). If \( a = (a_{j,l}) \in M_n(A) \), then put

\[
N_1(a) = \max_{1 \leq j,l \leq n} N_A(a_{j,l}) \\
N_{1,1}(a) = \max_{1 \leq j,l \leq n} \left( \sum_{j=1}^{n} N_A(a_{j,l}) \right) \\
N_{\infty,1}(a) = \max_{1 \leq j,l \leq n} \left( \sum_{i=1}^{n} N_A(a_{j,l}) \right)
\]

and

\[
N_\infty(a) = \max_{1 \leq j,l \leq n} N_A(a_{j,l}) \\
N_{\infty,1}(a) = \max_{1 \leq j,l \leq n} \left( \sum_{i=1}^{n} N_A(a_{j,l}) \right)
\]

It is easy to see that (3.3.1), (3.3.2), and (3.3.3) define seminorms on \( M_n(A) \), as a vector space over \( k \), and with respect to \( | \cdot | \) on \( k \). If \( N_A \) is a norm on \( A \), then (3.3.1), (3.3.2), and (3.3.3) are norms on \( M_n(A) \). Observe that

\[
N_\infty(a) \leq N_{1,\infty}(a) \leq n N_\infty(a)
\]

and

\[
N_\infty(a) \leq N_{\infty,1}(a) \leq n N_\infty(a)
\]

for every \( a \in M_n(A) \). In addition,

\[
N_\infty(a^t) = N_\infty(a)
\]

and

\[
N_{1,\infty}(a^t) = N_{\infty,1}(a)
\]

for every \( a \in M_n(A) \), where \( a^t \in M_n(A) \) is the transpose of \( a \), as before.

Suppose for the moment that \( A \) has a multiplicative identity element \( e \). Remember that the corresponding identity matrix \( I \in M_n(A) \) has diagonal entries equal to \( e \), and all other entries equal to 0. Thus

\[
N_\infty(I) = N_{1,\infty}(I) = N_{\infty,1}(I) = N_A(e).
\]

Let \( a, b \in M_n(A) \) be given, and let \( c = ab \) be their product, so that

\[
c_{j,r} = \sum_{l=1}^{n} a_{j,l} b_{l,r}
\]

for every \( j, r = 1, \ldots, n \). Observe that

\[
N_A(c_{j,r}) \leq \sum_{l=1}^{n} N_A(a_{j,l} b_{l,r}) \leq C \sum_{l=1}^{n} N_A(a_{j,l}) N_A(b_{l,r})
\]
for every $j, r = 1, \ldots, r$, using (3.2.1) in the second step. Thus

$$\sum_{j=1}^{n} N_A(c_{j,r}) \leq C \sum_{j=1}^{n} \sum_{l=1}^{n} N_A(a_{j,l}) N_A(b_{l,r})$$

(3.3.11)

$$= C \sum_{l=1}^{n} \sum_{j=1}^{n} N_A(a_{j,l}) N_A(b_{l,r})$$

$$\leq C N_{1,\infty}(a) \sum_{l=1}^{n} N_A(b_{l,r}) \leq C N_{1,\infty}(a) N_{1,\infty}(b)$$

for every $l = 1, \ldots, n$, so that

$$N_{1,\infty}(c) \leq C N_{1,\infty}(a) N_{1,\infty}(b).$$

(3.3.12)

Similarly,

$$\sum_{r=1}^{n} N_A(c_{j,r}) \leq C \sum_{r=1}^{n} \sum_{l=1}^{n} N_A(a_{j,l}) N_A(b_{l,r})$$

(3.3.13)

$$= C \sum_{l=1}^{n} \sum_{r=1}^{n} N_A(a_{j,l}) N_A(b_{l,r})$$

$$\leq C \sum_{l=1}^{n} N_A(a_{j,l}) N_{\infty,1}(b) \leq C N_{\infty,1}(a) N_{\infty,1}(b)$$

for every $j = 1, \ldots, n$, so that

$$N_{\infty,1}(c) \leq C N_{\infty,1}(a) N_{\infty,1}(b).$$

(3.3.14)

Suppose now that $N_A$ is a semi-ultranorm on $A$ with respect to $\cdot$ on $k$. This implies that $N_\infty$ is a semi-ultranorm on $M_n(A)$, as a vector space over $k$, and with respect to $\cdot$ on $k$. In this case, we have that

$$N_A(c_{j,r}) \leq \max_{1 \leq i \leq n} N_A(a_{j,i} b_{i,r}) \leq C \max_{1 \leq i \leq n} (N_A(a_{j,i}) N_A(b_{i,r}))$$

(3.3.15)

for every $j, r = 1, \ldots, n$, using (3.2.1) in the second step. It follows that

$$N_\infty(c) \leq C N_\infty(a) N_\infty(b).$$

(3.3.16)

### 3.4 Continuity of inverses

Let $k$ be a field with an absolute value function $|\cdot|$, and let $A$ be an associative algebra over $k$ with a multiplicative identity element $e$. Also let $N_A$ be a semi-norm on $A$ as a vector space over $k$, with respect to $|\cdot|$ on $k$, and suppose that $N_A$ satisfies the boundedness condition (3.2.1) with constant $C \geq 0$. If $a$ is an invertible element of $A$, then

$$N_A(e) \leq C N_A(a) N_A(a^{-1}).$$

(3.4.1)
Let \( x, \ y \) be invertible elements of \( A \), and observe that
\[
(x^{-1} - y^{-1}) = (x^{-1} (y y^{-1}) - (x^{-1} x) y^{-1} = x^{-1} (y - x) y^{-1}.
\]

It follows that
\[
N_A(x^{-1} - y^{-1}) \leq C^2 N_A(x^{-1}) N_A(y^{-1}) N_A(x - y).
\]

Hence
\[
N_A(y^{-1}) \leq N_A(x^{-1}) + N_A(x^{-1} - y^{-1}) \\
\leq N_A(x^{-1}) + C^2 N_A(x^{-1}) N_A(y^{-1}) N_A(x - y),
\]
so that
\[
(1 - C^2 N_A(x^{-1}) N_A(x - y)) N_A(y^{-1}) \leq N_A(x^{-1}).
\]

If
\[
C^2 N_A(x^{-1}) N_A(x - y) < 1,
\]
then we get that
\[
N_A(y^{-1}) \leq (1 - C^2 N_A(x^{-1}) N_A(x - y))^{-1} N_A(x^{-1}).
\]

Combining this with (3.4.3), we obtain that
\[
N_A(x^{-1} - y^{-1}) \leq C^2 (1 - C^2 N_A(x^{-1}))^{-1} N_A(x^{-1})^2 N_A(x - y)
\]
when (3.4.6) holds.

Suppose for the moment that \( N_A \) is a semi-ultranorm on \( A \), as a vector space over \( k \). Let us check that
\[
N_A(x^{-1}) = N_A(y^{-1})
\]
when (3.4.6) holds. Of course, this is trivial when \( N_A \equiv 0 \) on \( A \). Otherwise, if \( N_A \neq 0 \) on \( A \), then \( N_A(e) > 0 \), and hence \( N_A(y^{-1}) > 0 \), by (3.4.1). In this case, (3.4.6) implies that
\[
C^2 N_A(x^{-1}) N_A(y^{-1}) N_A(x - y) < N_A(y^{-1}).
\]

Combining this with (3.4.3), we get that
\[
N_A(x^{-1} - y^{-1}) < N_A(y^{-1}).
\]
This implies (3.4.9) in this situation, as in (1.8.7). It follows that
\[
N_A(x^{-1} - y^{-1}) \leq C^2 N_A(x^{-1})^2 N_A(x - y)
\]
when (3.4.6) holds, by (3.4.3) and (3.4.9).

Let us now take \( x = e \), for which there are some simplifications. If \( y \) is an invertible element of \( A \), then
\[
e^{-y^{-1}} = (y - e) y^{-1},
\]
3.5. BANACH ALGEBRAS

so that

\[ N_A(e - y^{-1}) \leq C N_A(y^{-1}) N_A(y - e). \]  

(3.4.14)

This implies that

\[ N_A(y^{-1}) \leq N_A(e) + N_A(e - y^{-1}) \leq N_A(e) + C N_A(y^{-1}) N_A(y - e), \]  

(3.4.15)

and hence

\[ (1 - C N_A(y - e)) N_A(y^{-1}) \leq N_A(e). \]  

If

\[ C N_A(y - e) < 1, \]  

(3.4.17)

then it follows that

\[ N_A(y^{-1}) \leq (1 - C N_A(y - e))^{-1} N_A(e). \]  

(3.4.18)

This implies that

\[ N_A(e - y^{-1}) \leq C (1 - C N_A(y - e))^{-1} N_A(e) N_A(y - e), \]  

(3.4.19)

when (3.4.17) holds, because of (3.4.14).

Suppose now that \( N_A \) is a semi-ultranorm on \( A \) again, and let us verify that

\[ N_A(y^{-1}) = N_A(e) \]  

(3.4.20)

when (3.4.17) holds. This is trivial when \( N_A \equiv 0 \) on \( A \), as before. Suppose instead that \( N_A \not\equiv 0 \) on \( A \), so that \( N_A(e) > 0 \), and thus \( N_A(y^{-1}) > 0 \). We can multiply both sides of (3.4.17) by \( N_A(y^{-1}) \), to get that

\[ C N_A(y^{-1}) N_A(y - e) < N_A(y^{-1}). \]  

(3.4.21)

This implies that

\[ N_A(e - y^{-1}) < N_A(y^{-1}), \]  

(3.4.22)

because of (3.4.14). This permits us to obtain (3.4.20) using (1.8.7), as before. Hence

\[ N_A(e - y^{-1}) \leq C N_A(e) N_A(y - e) \]  

(3.4.23)

when (3.4.17) holds, by (3.4.14) and (3.4.20).

3.5 Banach algebras

Let \( k \) be a field with an absolute value function \( | \cdot | \), and let \( A \) be an associative algebra over \( k \). Also let \( \| \cdot \| \) be a norm on \( A \) with respect to \( | \cdot | \) on \( k \) such that

\[ \| x y \| \leq C \| x \| \| y \| \]  

(3.5.1)

for some \( C \geq 0 \) and every \( x, y \in A \). In this section, we ask that \( A \) be complete with respect to the metric associated to \( \| \cdot \| \). Otherwise, one can pass to a
completion of \( A \), as usual. If (3.5.1) holds with \( C = 1 \), then \( A \) is said to be a Banach algebra with respect to \( \| \cdot \| \).

Let us suppose too that \( A \) has a nonzero multiplicative identity element \( e \), which is sometimes included in the definition of a Banach algebra. The condition

\[(3.5.2) \quad \| e \| = 1\]

is sometimes included in the definition of a Banach algebra as well.

Let \( a \in A \) be given, and remember that

\[(3.5.3) \quad (e - a) \sum_{j=0}^{n} a^j = \left( \sum_{j=0}^{n} a^j \right) (e - a) = e - a^{n+1}\]

for every nonnegative integer \( n \), as in (3.1.3). Observe that

\[(3.5.4) \quad \|a^j\| \leq C^{j-1} \|a\|^j\]

for every positive integer \( j \), by (3.5.1). Suppose that

\[(3.5.5) \quad C \|a\| < 1,\]

so that

\[(3.5.6) \quad \lim_{j \to \infty} \|a^j\| \to 0,\]

by (3.5.4). We also get that

\[(3.5.7) \quad \sum_{j=0}^{\infty} \|a^j\|\]

converges as an infinite series of nonnegative real numbers, because

\[(3.5.8) \quad \sum_{j=0}^{\infty} C^j \|a\|^j\]

is a convergent geometric series. This means that \( \sum_{j=0}^{\infty} a^j \) converges absolutely with respect to \( \| \cdot \| \), and hence that \( \sum_{j=0}^{\infty} a^j \) converges in \( A \), because \( A \) is complete with respect to the metric associated to \( \| \cdot \| \). The value of this sum satisfies

\[(3.5.9) \quad (e - a) \sum_{j=0}^{\infty} a^j = \left( \sum_{j=0}^{\infty} a^j \right) (a - e) = e,\]

by taking the limit as \( n \to \infty \) in (3.5.3). Thus \( e - a \) is invertible in \( A \), with

\[(3.5.10) \quad (e - a)^{-1} = \sum_{j=0}^{\infty} a^j.\]

Let \( x \) be an invertible element of \( A \), and let \( y \) be another element of \( A \). Observe that

\[(3.5.11) \quad y = x - (x - y) = x (e - x^{-1} (x - y)).\]
Suppose that
\[ C^2 \|x^{-1}\| \|x - y\| < 1, \]
so that
\[ C \|x^{-1}(x - y)\| \leq C^2 \|x^{-1}\| \|x - y\| < 1. \]
This implies that \( e - x^{-1}(x - y) \) is invertible in \( A \), as in the preceding paragraph. It follows that \( y \) is invertible in \( A \), by (3.5.11).

Let \( a \) be an element of \( A \) again, and let \( j_0 \) be a positive integer. If \( e - a^{j_0} \) is invertible in \( A \), then \( e - a \) is invertible in \( A \), because of (3.5.3), with \( n = j_0 - 1 \). In particular, this holds when
\[ C \|a^{j_0}\| < 1; \]
as before. Alternatively, one can use (3.5.1) to estimate \( \|a^{j_0}l + r\| \) in terms of \( C^l \|a^{j_0}\|^l \) when \( l \geq 1 \) and \( 0 \leq r < j_0 \), to get that (3.5.6) holds and that (3.5.7) converges when (3.5.14) holds. This implies that \( \sum_{j=0}^{\infty} a^j \) converges in \( A \) and satisfies (3.5.9) when (3.5.14) holds, as before.

## 3.6 Invertible linear mappings

Let \( k \) be a field with an absolute value function \( |\cdot| \), and let \( V, W \) be vector spaces over \( k \) with seminorms \( N_V, N_W \), respectively, with respect to \( |\cdot| \) on \( k \). If \( T \) is a one-to-one linear mapping from \( V \) onto \( W \), then the corresponding inverse mapping \( T^{-1} \) is a linear mapping from \( W \) onto \( V \). As usual, \( T^{-1} \) is bounded with respect to \( N_W, N_V \) if there is a nonnegative real number \( C \) such that
\[ N_V(T^{-1}(w)) \leq C N_W(w) \]
for every \( w \in W \). This is the same as saying that
\[ N_V(v) \leq C N_W(T(v)) \]
for every \( v \in V \).

Now let \( T \) be a linear mapping from \( V \) into \( W \), and suppose that (3.6.2) holds for some \( C \geq 0 \). If \( N_V \) is a norm on \( V \), then it follows that \( T \) is injective on \( V \). Let \( T_1 \) be another linear mapping from \( V \) into \( W \), and observe that
\[ N_V(v) \leq C N_W(T_1(v)) + C N_W(T_1(v) - T(v)) \]
for every \( v \in V \). If \( T_1 - T \) is bounded as a linear mapping from \( V \) into \( W \), then we get that
\[ N_V(v) \leq C N_W(T_1(v)) + C \|T_1 - T\|_{op,VW} N_V(v) \]
for every \( v \in V \), where the operator seminorm \( \|\cdot\|_{op,VW} \) is as defined in Section 1.9. Thus
\[ (1 - C \|T_1 - T\|_{op,VW}) N_V(v) \leq C N_W(T_1(v)) \]
for every $v \in V$. If
\[(3.6.6) \quad C \|T_1 - T\|_{op,VW} < 1,\]
then it follows that
\[(3.6.7) \quad N_V(v) \leq C (1 - C \|T_1 - T\|_{op,VW})^{-1} N_W(T_1(v))\]
for every $v \in V$.

Suppose that $N_W$ is a semi-ultranorm on $W$ with respect to $|\cdot|$ on $k$. In this case, we have that
\[(3.6.8) \quad N_V(v) \leq C \max(N_W(T_1(v)), N_W(T_1(v) - T(v)))\]
for every $v \in V$, by (3.6.2). Suppose for the moment that
\[(3.6.9) \quad C N_W(T_1(v) - T(v)) < N_V(v)\]
for every $v \in V$ with $N_V(v) > 0$. It follows that
\[(3.6.10) \quad N_V(v) \leq C N_W(T_1(v))\]
for every $v \in V$ with $N_V(v) > 0$. Of course, (3.6.10) holds trivially when $N_V(v) = 0$, so that (3.6.10) holds for all $v \in V$. If $T_1 - T$ is bounded as a linear mapping from $V$ into $W$, then
\[(3.6.11) \quad N_W(T_1(v) - T(v)) \leq \|T_1 - T\|_{op,VW} N_V(v)\]
for every $v \in V$. If (3.6.6) holds, then we get that (3.6.9) holds when $N_V(v) > 0$.

A bounded linear mapping $T$ from $V$ onto $W$ is said to be invertible as a bounded linear mapping if $T$ is a one-to-one mapping from $V$ onto $W$ whose inverse $T^{-1}$ is bounded with respect to $N_W$, $N_V$. The bounded linear mappings from $V$ onto itself with bounded inverses are the same as the invertible elements of $\mathcal{B}L(V)$, as an algebra with respect to composition of mappings.

Suppose that $N_V$ and $N_W$ are norms on $V$ and $W$, respectively, and that $T$ is a bounded linear mapping from $V$ into $W$ that satisfies (3.6.2) for some $C \geq 0$. If $V$ is complete with respect to the metric associated to $N_V$, then it is easy to see that the image $T(V)$ of $V$ under $T$ is complete with respect to the restriction of the metric on $W$ associated to $N_W$ to $T(V)$. This implies that $T(V)$ is a closed set in $W$ with respect to the metric associated to $N_W$, by a standard argument.

### 3.7 Isometric linear mappings

Let $k$ be a field with an absolute value function $|\cdot|$ again, and let $V$, $W$ be vector spaces over $k$ with seminorms $N_V$, $N_W$, respectively, with respect to $|\cdot|$ on $k$. A linear mapping $T$ from $V$ into $W$ is said to be an isometry with respect to $N_V$ and $N_W$ if
\[(3.7.1) \quad N_W(T(v)) = N_V(v)\]
for every $v \in V$. Of course, this is the same as saying that

\[(3.7.2) \quad N_W(T(v)) \leq N_V(v)\]

and

\[(3.7.3) \quad N_V(v) \leq N_W(T(v))\]

for every $v \in V$. The first condition (3.7.2) means that $T$ is a bounded linear mapping from $V$ into $W$ with respect to $N_V$ and $N_W$, with

\[(3.7.4) \quad \|T\|_{op;VW} \leq 1,\]

where the operator seminorm is as defined in Section 1.9. The second condition (3.7.3) is the same as (3.6.2), with $C = 1$.

Let $Z$ be another vector space over $k$ with a seminorm $N_Z$ with respect to $\| \cdot \|$ on $k$. Also let $T_1$ be an isometric linear mapping from $V$ into $W$ with respect to $N_V$ and $N_W$, and let $T_2$ be an isometric linear mapping from $W$ into $Z$ with respect to $N_W$ and $N_Z$. Observe that

\[(3.7.5) \quad N_Z((T_2 \circ T_1)(v)) = N_Z(T_2(T_1(v))) = N_W(T_1(v)) = N_V(v)\]

for every $v \in V$, so that $T_2 \circ T_1$ is an isometric linear mapping from $V$ into $Z$.

If $T$ is a one-to-one linear mapping from $V$ onto $W$, then (3.7.3) is equivalent to saying that $T^{-1}$ is a bounded linear mapping from $W$ into $V$ with respect to $N_W$ and $N_V$, with

\[(3.7.6) \quad \|T^{-1}\|_{op;WV} \leq 1,\]

as in the previous section. Thus $T$ is an isometric linear mapping if and only $T$ and $T^{-1}$ are bounded linear mappings that satisfy (3.7.4) and (3.7.6). In particular, $T$ is an isometric linear mapping if and only if $T^{-1}$ is an isometric linear mapping. Of course, the identity mapping on $V$ is an isometric linear mapping from $V$ onto itself with respect to $N_V$. The collection of one-to-one isometric linear mappings from $V$ onto itself is a group with respect to composition of mappings.

Suppose now that $N_W$ is a semi-ultranorm on $W$ with respect to $\| \cdot \|$ on $k$. Let $T$ be a linear mapping from $V$ into $W$ that satisfies (3.7.3) for every $v \in V$. Let $T_1$ be another linear mapping from $V$ into $W$ such that

\[(3.7.7) \quad N_W(T_1(v) - T(v)) < N_V(v)\]

for every $v \in V$ with $N_V(v) > 0$. Under these conditions, we have that

\[(3.7.8) \quad N_V(v) \leq N_W(T_1(v))\]

for every $v \in V$, as in (3.6.10), with $C = 1$.

Let $T$ be an isometric linear mapping from $V$ into $W$, and let $T_1$ be a bounded linear mapping from $V$ into $W$. If

\[(3.7.9) \quad \|T_1 - T\|_{op,VW} \leq 1,\]
then
\[\|T_1\|_{\text{op},VW} \leq 1,\]  
(3.7.10)

because of (3.7.4) and the hypothesis that \(N_W\) be a semi-ultranorm on \(W\). If
\[\|T_1 - T\|_{\text{op},VW} < 1,\]  
(3.7.11)

then (3.7.7) holds for every \(v \in V\) with \(N_V(v) > 0\), so that \(T_1\) satisfies (3.7.8), as before. This shows that \(T_1\) is also an isometric linear mapping from \(V\) into \(W\) when (3.7.11) holds.

3.8 Hilbert space isometries

Let \((V,\langle \cdot,\cdot \rangle_V)\) and \((W,\langle \cdot,\cdot \rangle_W)\) be inner product spaces, both real or both complex, and let \(\|\cdot\|_V\) and \(\|\cdot\|_W\) be the corresponding norms on \(V\) and \(W\), respectively, as in Section 1.11. If a linear mapping \(T\) from \(V\) into \(W\) satisfies
\[\langle T(u),T(v) \rangle_W = \langle u,v \rangle_V\]  
(3.8.1)

for every \(u,v \in V\), then it is easy to see that \(T\) is an isometry with respect to \(\|\cdot\|_V\) and \(\|\cdot\|_W\), by taking \(u = v\). Conversely, if \(T\) is an isometric linear mapping from \(V\) into \(W\) with respect to \(\|\cdot\|_V\) and \(\|\cdot\|_W\), then one can check that \(T\) satisfies (3.8.1), using polarization identities. An isometric linear mapping from \(V\) onto \(W\) is also known as an orthogonal transformation in the real case, and a unitary transformation in the complex case. The orthogonal or unitary transformations from \(V\) onto itself form a group with respect to composition of mappings.

Suppose that \(V\) and \(W\) are Hilbert spaces, and that \(T\) is a bounded linear mapping from \(V\) into \(W\). Let \(T^*\) be the corresponding adjoint mapping from \(W\) into \(V\), as in Section 1.11. Observe that
\[\langle T(u),T(v) \rangle_W = \langle u,T^*(T(v)) \rangle_V\]  
(3.8.2)

for every \(u,v \in V\). Thus \(T\) is an isometric linear mapping from \(V\) into \(W\) if and only if
\[T^* \circ T = I_V,\]  
(3.8.3)

where \(I_V\) is the identity mapping on \(V\). This is the same as saying that
\[T^* = T^{-1}\]  
(3.8.4)

when \(T\) maps \(V\) onto \(W\).

If \(T\) is any bounded linear mapping from \(V\) into \(W\), then it is well known that
\[\|T^* \circ T\|_{\text{op},VW} = \|T\|^2_{\text{op},W},\]  
(3.8.5)

where these operator norms are taken with respect to \(\|\cdot\|_V\) and \(\|\cdot\|_W\), as appropriate. More precisely,
\[\|T^* \circ T\|_{\text{op},V} \leq \|T\|_{\text{op},W} \|T^*\|_{\text{op},V} \|T\|^2_{\text{op},W},\]  
(3.8.6)
3.9. PRESERVING BILINEAR FORMS

using (1.11.15) in the second step. We also have that

\[
\|T(v)\|_W^2 = \langle v, T^*(T(v)) \rangle_V \leq \|v\|_V \|T^*(T(v))\|_V
\]

\[
\leq \|T^* \circ T\|_{op,VV} \|v\|_V^2
\]

(3.8.7)

for every \(v \in V\), by taking \(u = v\) in (3.8.2) in the first step, and using the Cauchy–Schwarz inequality in the second step. This implies that

\[
\|T\|_{op,VW}^2 \leq \|T^* \circ T\|_{op,VV}, \quad (3.8.8)
\]

as desired.

Observe that \(w \in W\) satisfies \(T^*(w) = 0\) if and only if

\[
\langle T^*(v), w \rangle_W = 0 \quad (3.8.9)
\]

for every \(v \in V\), by the definition of \(T^*\). If the image \(T(V)\) of \(V\) under \(T\) is dense in \(W\) with respect to the metric associated to \(\|\cdot\|_W\), then it follows that \(T^*(w) = 0\). However, if \(T(V)\) is not dense in \(W\), then there is a \(w \in W\) such that \(w \neq 0\) and (3.8.9) holds for every \(v \in V\), by standard results about Hilbert spaces. Thus \(T(V)\) is dense in \(W\) if and only if the kernel of \(T^*\) is trivial.

Suppose that

\[
\|v\|_V \leq C \|T(v)\|_W \quad (3.8.10)
\]

for some nonnegative real number \(C\) and every \(v \in V\). This implies that \(T(V)\) is a closed set in \(W\) with respect to the metric associated to \(\|\cdot\|_W\), as in Section 3.6, because \(V\) is complete, by hypothesis. If \(T(V)\) is dense in \(W\), then it follows that \(T(V) = W\).

3.9 Preserving bilinear forms

Let \(k\) be a commutative ring with a multiplicative identity element, and let \(A_1\), \(A_2\), and \(C\) be modules over \(k\). Also let \(\beta_1, \beta_2\) be bilinear mappings from \(A_1 \times A_1\) and \(A_2 \times A_2\) into \(C\), respectively, as in Section 2.12. Let us say that a module homomorphism \(\phi\) from \(A_1\) into \(A_2\) preserves these bilinear mappings if

\[
\beta_2(\phi(a_1), \phi(b_1)) = \beta_1(a_1, b_1) \quad (3.9.1)
\]

for every \(a_1, b_1 \in A_1\). If \(\phi\) is a one-to-one mapping from \(A_1\) onto \(A_2\), then (3.9.1) is equivalent to asking that

\[
\beta_2(\phi^{-1}(a_2), \phi^{-1}(b_2)) = \beta_2(a_2, b_2) \quad (3.9.2)
\]

for every \(a_2, b_2 \in A_2\). Of course, this is the same as saying that \(\phi^{-1}\) preserves \(\beta_2, \beta_1\).

Let \(A_3\) be another module over \(k\), and let \(\beta_3\) be a bilinear mapping from \(A_3 \times A_3\) into \(C\). Suppose that \(\phi_1\) is a module homomorphism from \(A_1\) into \(A_2\)
that preserves $\beta_1$, $\beta_2$, and that $\phi_2$ is a module homomorphism from $A_2$ into $A_3$ that preserves $\beta_2$, $\beta_3$. This implies that

\[(3.9.3) \quad \beta_3(\phi_2(\phi_1(a_1)), \phi_2(\phi_1(b_1))) = \beta_2(\phi_1(a_1), \phi_1(b_1)) = \beta_1(a_1, b_1)\]

for every $a_1, b_1 \in A_1$, so that $\phi_2 \circ \phi_1$ is a module homomorphism from $A_1$ into $A_3$ that preserves $\beta_1$, $\beta_3$.

Let $\phi$ be a module homomorphism from $A_1$ into $A_2$, and suppose that

\[(3.9.4) \quad \beta_2(\phi(a), \phi(a)) = \beta_1(a, a)\]

for every $a \in A_1$. If $a, b \in A_1$, then

\[(3.9.5) \quad \beta_1(a + b, a + b) = \beta_1(a, a) + \beta_1(a, b) + \beta_1(b, a) + \beta_1(b, b)\]

and

\[(3.9.6) \quad \beta_2(\phi(a + b), \phi(a + b)) = \beta_2(\phi(a) + \phi(b), \phi(a) + \phi(b)) = \beta_2(\phi(a), \phi(a)) + \beta_2(\phi(b), \phi(b)) + \beta_2(\phi(a), \phi(b)) + \beta_2(\phi(b), \phi(a)).\]

It follows that

\[(3.9.7) \quad \beta_2(\phi(a), \phi(b)) + \beta_2(\phi(b), \phi(a)) = \beta_1(a, b) + \beta_1(b, a)\]

for every $a, b \in A_1$. If $\beta_1$ and $\beta_2$ are symmetric bilinear mappings, and if $1 + 1$ is invertible in $k$, then we get that $\phi$ preserves $\beta_1$, $\beta_2$. Of course, (3.9.4) holds when $\phi$ preserves $\beta_1$, $\beta_2$.

Let $A$ be a module over $k$, and remember that the space $\text{Hom}_k(A, A)$ of module homomorphisms from $A$ into itself is an associative algebra over $k$ with respect to composition of mappings. If $\phi$ is a one-to-one module homomorphism from $A$ onto itself, then $\phi^{-1}$ is a module homomorphism from $A$ into itself as well. In this case, $\phi$ may be called an module automorphism of $A$. The module automorphisms of $A$ are the same as the invertible elements of $\text{Hom}_k(A, A)$, and form a group with respect to composition of mappings.

A bilinear mapping $\beta$ from $A \times A$ into $C$ is said to be invariant under an module automorphism $\phi$ on $A$ if

\[(3.9.8) \quad \beta(\phi(a), \phi(b)) = \beta(a, b)\]

for every $a, b \in A$, which is the same as saying that $\phi$ preserves $\beta$ as a bilinear mapping from $A \times A$ into $C$ for both the domain and range. The identity mapping on $A$ obviously has this property. The collection of module automorphisms of $A$ that preserve $\beta$ is a subgroup of the group of all module automorphisms of $A$. 
3.10 Preserving nondegenerate bilinear forms

Let \( k \) be a field, let \( V_1, V_2 \) be vector spaces over \( k \), and let \( b_1, b_2 \) be bilinear forms on \( V_1, V_2 \), respectively. Suppose that \( V_1 \) has finite dimension, and that \( b_1 \) is nondegenerate on \( V_1 \), as in Section 2.14. Let \( T \) be a linear mapping from \( V_1 \) into \( V_2 \). If \( w_2 \in V_2 \), then \( b_2(T(v_1), w_2) \) defines a linear functional on \( V_1 \), as a function of \( v_1 \). This implies that there is a unique element \( T^*(w_2) \) of \( V_1 \) such that

\[
(3.10.1) \quad b_2(T(v_1), w_2) = b_1(v_1, T^*(w_2))
\]

for every \( v_1 \in V_1 \). This defines a linear mapping \( T^* \) from \( V_2 \) into \( V_1 \), which is the adjoint of \( T \) with respect to \( b_1, b_2 \). The mapping \( T \mapsto T^* \) is linear as a mapping from the space \( \mathcal{L}(V_1, V_2) \) of linear mappings from \( V_1 \) into \( V_2 \) into the corresponding space \( \mathcal{L}(V_2, V_1) \).

Let \( V_3 \) be another vector space over \( k \) with a bilinear form \( b_3 \), and suppose that \( V_3 \) also has finite dimension, and that \( b_3 \) is nondegenerate on \( V_3 \). If \( T_1 \) is a linear mapping from \( V_1 \) into \( V_2 \), and \( T_2 \) is a linear mapping from \( V_2 \) into \( V_3 \), then their adjoints \( T_1^* \) and \( T_2^* \) can be defined as in the preceding paragraph. Similarly, \( T_2 \circ T_1 \) is a linear mapping from \( V_1 \) into \( V_3 \), whose adjoint can be defined as in the previous paragraph as well. Observe that

\[
(3.10.2) \quad b_3(T_2 \circ T_1)(v_1), w_3) = b_3(T_2(T_1(v_1)), w_3) = b_2(T_1(v_1), T_2^*(w_3)) = b_1(v_1, T_1^*(T_2^*(w_3))) = b_1(v_1, (T_1^* \circ T_2^*)(w_3))
\]

for every \( v_1 \in V_1 \) and \( w_3 \in V_3 \). This implies that

\[
(3.10.3) \quad (T_2 \circ T_1)^* = T_1^* \circ T_2^*
\]

as linear mappings from \( V_3 \) into \( V_1 \).

Let us continue to suppose for the moment that \( V_2 \) has finite dimension, and that \( b_2 \) is nondegenerate on \( V_2 \). If \( T \) is a linear mapping from \( V_1 \) into \( V_2 \) and \( T^* = 0 \) on \( V_2 \), then \( T = 0 \) on \( V_1 \), because of (3.10.1) and the nondegeneracy of \( b_2 \) on \( V_2 \). This implies that \( T \mapsto T^* \) is a one-to-one mapping from \( \mathcal{L}(V_1, V_2) \) onto \( \mathcal{L}(V_2, V_1) \), because \( \mathcal{L}(V_1, V_2) \) and \( \mathcal{L}(V_2, V_1) \) are finite-dimensional vector spaces over \( k \) with the same dimension. Let \( T \) be a linear mapping from \( V_1 \) into \( V_2 \) again, and note that the adjoint \( (T^*)^* \) of the adjoint \( T^* \) of \( T \) can be defined as a linear mapping from \( V_1 \) into \( V_2 \) in the same way in this situation. If \( b_1 \) and \( b_2 \) are symmetric on \( V_1 \) and \( V_2 \), respectively, then

\[
(3.10.4) \quad b_2(T(v_1), w_2) = b_1(T^*(w_2), v_1) = b_2(w_2, (T^*)^*(v_1)) = b_2((T^*)^*(v_1), w_2)
\]

for every \( v_1 \in V_1 \) and \( w_2 \in V_2 \). Similarly, if \( b_1 \) and \( b_2 \) are antisymmetric on \( V_1 \) and \( V_2 \), respectively, then

\[
(3.10.5) \quad b_2(T(v_1), w_2) = -b_1(T^*(w_2), v_1) = -b_2(w_2, (T^*)^*(v_1)) = b_2((T^*)^*(v_1), w_2)
\]
CHAPTER 3. SUBMULTIPLICATIVITY AND INVERTIBILITY

for every \( v_1 \in V_1 \) and \( w_2 \in V_2 \). In both cases, it follows that

\[(T^*)^* = T.\]  

(3.10.6)

If one of \( b_1 \) and \( b_2 \) is symmetric, and the other is antisymmetric, then

\[(T^*)^* = -T,\]

(3.10.7)

by the analogous argument.

If \( T \) is any linear mapping from \( V_1 \) into \( V_2 \), then

\[b_2(T(v_1), T(w_1)) = b_1(v_1, T^*(T(w_1)))\]

(3.10.8)

for every \( v_1, w_1 \in V_1 \), by the definition of \( T^* \). Thus \( T \) preserves \( b_1, b_2 \), as in the previous section, if and only if

\[T^* \circ T = I_{V_1},\]

(3.10.9)

where \( I_{V_1} \) is the identity mapping on \( V_1 \). In particular, this implies that \( T \) is injective, which could also be obtained more directly from the nondegeneracy of \( b_1 \) on \( V_1 \). If \( V_2 \) has the same dimension as \( V_1 \), then it follows that \( T \) maps \( V_1 \) onto \( V_2 \). In this case, we get that

\[T^* = T^{-1}.\]  

(3.10.10)

3.11 Preserving sesquilinear forms

Let \( V_1, V_2 \) be vector spaces over the complex numbers, and let \( b_1, b_2 \) be sesquilinear forms on \( V_1, V_2 \), respectively. Let us say that a linear mapping \( T \) from \( V_1 \) into \( V_2 \) preserves \( b_1, b_2 \) if

\[b_2(T(v_1), T(w_1)) = b_1(v_1, w_1)\]

(3.11.1)

for every \( v_1, w_1 \in V_1 \). If \( T \) is a one-to-one linear mapping from \( V_1 \) onto \( V_2 \), then this is the same as saying that

\[b_1(T^{-1}(v_2), T^{-1}(w_2)) = b_2(v_2, w_2)\]

(3.11.2)

for every \( v_2, w_2 \in V_2 \), which means that \( T^{-1} \) preserves \( b_2, b_1 \).

Let \( V_3 \) be another complex vector space with a sesquilinear form \( b_3 \). If \( T_1 \) is a linear mapping from \( V_1 \) into \( V_2 \) that preserves \( b_1, b_2 \), and \( T_2 \) is a linear mapping from \( V_2 \) into \( V_3 \) that preserves \( b_2, b_3 \), then it is easy to see that their composition \( T_2 \circ T_1 \) preserves \( b_1, b_3 \).

Let \( T \) be a linear mapping from \( V_1 \) into \( V_2 \) that satisfies

\[b_2(T(u_1), T(u_1)) = b_1(u_1, u_1)\]

(3.11.3)

for every \( u_1 \in V \). One can check that \( T \) satisfies (3.11.1) for every \( v_1, w_1 \in V_1 \), by applying (3.11.3) to \( u_1 = v_1 + w_1 \) and to \( u_1 = v_1 + i w_1 \). Of course, (3.11.1) implies (3.11.3), by taking \( v_1, w_1 = u_1 \).
3.11. PRESERVING SESQUILINEAR FORMS

Let $V$ be a complex vector space, and let $T$ be a one-to-one linear mapping from $V$ onto itself. A sesquilinear form $b$ on $V$ is said to be invariant under $T$ if

$$b(T(v), T(w)) = b(v, w)$$

(3.11.4)

for every $v, w \in V$, which is to say that $T$ preserves $b$ as a sesquilinear form on both the domain and range. The collection of one-to-one linear mappings from $V$ onto itself that preserve $b$ is a group with respect to composition.

Let $V_1$, $V_2$ be complex vector spaces again, and let $b_1$, $b_2$ be sesquilinear forms on them, respectively. Let us suppose for the rest of the section that $V_1$ has finite dimension, and that $b_1$ is nondegenerate on $V_1$, as in Section 2.15.

Let $T$ be a linear mapping from $V_1$ into $V_2$, and let $w_2 \in V_2$ be given. Thus $b_2(T(v_1), w_2)$ is a linear functional on $V_1$, as a function of $v_1$, so that there is a unique element $T^*(w_2)$ of $V_1$ such that

$$b_2(T(v_1), w_2) = b_1(v_1, T^*(w_2))$$

(3.11.5)

for every $v_1 \in V_1$. One can check that $T^*$ is a linear mapping from $V_2$ into $V_1$, and that the mapping from $T$ to its adjoint $T^*$ is conjugate-linear as a mapping from $L(V_1, V_2)$ into $L(V_2, V_1)$.

Let $V_3$ be another complex vector space with a sesquilinear form $b_3$, and suppose that $V_2$ has finite dimension, and that $b_2$ is nondegenerate on $b_2$. If $T_1$ is a linear mapping from $V_1$ into $V_2$, and $T_2$ is a linear mapping from $V_2$ into $V_3$, then $T_2 \circ T_1$ is a linear mapping from $V_1$ into $V_3$, and the adjoints of $T_1$, $T_2$, and $T_3$ can be defined as in the preceding paragraph. Under these conditions, one can verify that

$$T_2 \circ T_1 = T_1^* \circ T_2^*$$

(3.11.6)

as linear mappings from $V_3$ into $V_1$.

Let us continue to ask for the moment that $V_2$ have finite dimension, and that $b_2$ be nondegenerate on $V_2$. If $T$ is a linear mapping from $V_1$ into $V_2$ such that $T^* = 0$ on $V_2$, then $T = 0$ on $V_1$, because of (3.11.5) and the nondegeneracy of $b_2$ on $V_2$. It follows that $T \mapsto T^*$ is a one-to-one mapping from $L(V_1, V_2)$ onto $L(V_2, V_1)$, because $L(V_1, V_2)$ and $L(V_2, V_1)$ are finite-dimensional vector spaces over $\mathbb{C}$ with the same dimension. Let $T$ be a linear mapping from $V_1$ into $V_2$ again, so that the adjoint $(T^*)^*$ of $T^*$ is defined as a linear mapping from $V_1$ into $V_2$. Suppose that $b_1$ and $b_2$ are Hermitian-symmetric on $V_1$ and $V_2$, respectively. Under these conditions, we have that

$$b_2(T(v_1), w_2) = b_1(T^*(w_2), v_1) = b_2(w_2, (T^*)^*(v_1)) = b_2((T^*)^*(v_1), w_2)$$

(3.11.7)

for every $v_1 \in V_1$ and $w_2 \in V_2$. This implies that

$$(T^*)^* = T$$

(3.11.8)

in this situation.
As before,
\[(3.11.9)\]
\[b_2(T(v_1), T(w_1)) = b_1(v_1, T^*(T(w_1)))\]
for every linear mapping \(T\) from \(V_1\) into \(V_2\) and \(v_1, w_1 \in V_1\). This implies that \(T\) preserves \(b_1, b_2\) if and only if
\[(3.11.10)\]
\[T^* \circ T = I_{V_1},\]
because \(b_1\) is nondegenerate on \(V_1\). Note that \(T\) is injective in this case. If \(T\) maps \(V_1\) onto \(V_2\), then (3.11.10) is the same as saying that
\[(3.11.11)\]
\[T^* = T^{-1}.\]
Of course, surjectivity of \(T\) follows from injectivity when \(V_2\) has the same dimension as \(V_1\).

### 3.12 Bilinear forms and matrices

Let \(k\) be a commutative ring with a multiplicative identity element, and let \(n\) be a positive integer. The space \(k^n\) of \(n\)-tuples of elements of \(k\) is a (free) module over \(k\) with respect to coordinatewise addition and scalar multiplication. Let \(C\) be another module over \(k\), and let \((\beta_{j,l})\) be an \(n \times n\) matrix with entries in \(C\).

Put
\[(3.12.1)\]
\[\beta(x, y) = \sum_{j=1}^{n} \sum_{l=1}^{n} \beta_{j,l} x_l y_j\]
for every \(x, y \in k^n\), where the terms of the sum are defined using multiplication on \(k\) and scalar multiplication on \(C\). This defines a mapping from \(k^n \times k^n\) into \(C\) that is bilinear over \(k\), and it is easy to see that every bilinear mapping from \(k^n \times k^n\) into \(C\) can be expressed as (3.12.1) in a unique way.

Observe that (3.12.1) is symmetric as a bilinear mapping from \(k^n \times k^n\) into \(C\) if and only if \((\beta_{j,l})\) is symmetric as a matrix, which is to say that
\[(3.12.2)\]
\[\beta_{i,j} = \beta_{j,i}\]
for every \(j, l = 1, \ldots, n\). Similarly, (3.12.1) is antisymmetric as a bilinear mapping from \(k^n \times k^n\) into \(C\) if and only if \((\beta_{j,l})\) is antisymmetric as a matrix, in the sense that
\[(3.12.3)\]
\[\beta_{i,j} = -\beta_{j,i}\]
for every \(j, l = 1, \ldots, n\). Remember that (3.12.1) is antisymmetric as a bilinear mapping from \(k^n \times k^n\) into \(C\) when
\[(3.12.4)\]
\[\beta(x, x) = 0\]
for every \(x \in k^n\), as in Section 2.1. In this situation, one can check that (3.12.4) holds for every \(x \in k^n\) if and only if \((\beta_{j,l})\) is antisymmetric and
\[(3.12.5)\]
\[\beta_{j,j} = 0\]
for every \( j = 1, \ldots, n \). If \( 1 + 1 \) is invertible in \( k \), then (3.12.3) implies (3.12.5), by taking \( j = l \).

Let \( a = (a_{j,l}) \) be an \( n \times n \) matrix with entries in \( k \). If \( x \in k^n \), then let \( T_a(x) \) be the element of \( k^n \) whose \( j \)th coordinate is given by

\[
(T_a(x))_j = \sum_{l=1}^{n} a_{j,l} x_l
\]

for each \( j = 1, \ldots, n \). This defines a module homomorphism from \( k^n \) into itself, and every module homomorphism from \( k^n \) into itself corresponds to a unique \( a \in M_n(k) \) in this way. More precisely, \( a \mapsto T_a \) is an isomorphism from \( M_n(k) \) as an algebra over \( k \) with respect to matrix multiplication onto the algebra \( \text{Hom}_k(k^n, k^n) \) of module homomorphisms from \( k^n \) into itself with respect to composition of mappings. This corresponds to some of the remarks in Section 2.10 when \( k \) is a field.

Let \( a \in M_n(k) \) be given again, and observe that

\[
\beta(T_a(x), y) = \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{r=1}^{n} \beta_{r,j} a_{j,l} x_l y_r
\]

for every \( x, y \in k^n \), where the terms of the sum are again defined using multiplication on \( k \) and scalar multiplication on \( C \). Similarly,

\[
\beta(x, T_a(y)) = \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{r=1}^{n} \beta_{j,l} a_{j,r} x_l y_r
\]

for every \( x, y \in k^n \). It follows that \( T_a \) is symmetric with respect to \( \beta \), as in Section 2.12, if and only if

\[
\sum_{j=1}^{n} \beta_{r,j} a_{j,l} = \sum_{j=1}^{n} \beta_{j,l} a_{j,r}
\]

for every \( l, r = 1, \ldots, n \). Similarly, \( T_a \) is antisymmetric with respect to \( \beta \) if and only if

\[
\sum_{j=1}^{n} \beta_{r,j} a_{j,l} = -\sum_{j=1}^{n} \beta_{j,l} a_{j,r}
\]

for every \( l, r = 1, \ldots, n \).

We also have that

\[
\beta(T_a(x), T_a(y)) = \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{m=1}^{n} \sum_{r=1}^{n} \beta_{m,j} a_{j,l} a_{m,r} x_l y_r
\]

for every \( x, y \in k^n \), where the terms of the sum are defined using multiplication on \( k \) and scalar multiplication on \( C \). Thus \( T_a \) preserves \( \beta \), as in Section 3.9, if and only if

\[
\sum_{j=1}^{n} \sum_{m=1}^{n} \beta_{m,j} a_{j,l} a_{m,r} = \beta_{r,l}
\]
for every $l, r = 1, \ldots, n$. Note that $T_a$ is a module automorphism of $k^n$ exactly when $a$ is invertible in $M_n(k)$.

Let $b \in M_n(k)$ be given, so that $T_b$ can be defined as before, and

$$\beta(x, T_b(y)) = \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{r=1}^{n} \beta_{j,l} b_{j,r} x_l y_r$$

(3.12.13) for every $x, y \in k^n$, as in (3.12.8). Comparing this with (3.12.7), we get that

$$\beta(T_a(x), y) = \beta(x, T_b(y))$$

(3.12.14) for every $x, y \in k^n$ if and only if

$$\sum_{j=1}^{n} \beta_{r,j} a_{j,l} = \sum_{j=1}^{n} \beta_{j,l} b_{j,r}$$

(3.12.15) for every $l, r = 1, \ldots, n$.

The product of an $n \times n$ matrix with entries in $C$ and an $n \times n$ matrix with entries in $k$, in either order, can be defined as an $n \times n$ matrix with entries in $C$ in the usual way. Let us also use $\beta$ to denote $(\beta_{j,l})$, as an element of $M_n(C)$. Thus (3.12.9) is the same as saying that

$$\beta a = a^t \beta$$

(3.12.16) as elements of $M_n(C)$, where $a^t$ is the transpose of $a$, as in Section 2.8. Similarly, (3.12.10) is the same as saying that

$$\beta a = -a^t \beta$$

(3.12.17) as elements of $M_n(C)$. We can reexpress (3.12.12) as

$$a^t \beta a = \beta,$$

(3.12.18) and (3.12.15) as

$$\beta a = b^t \beta.$$  

(3.12.19)

Let us now take $C = k$, as a module over itself with respect to multiplication on $k$. If $\beta$ is invertible in $M_n(k)$, then (3.12.19) is the same as saying that

$$b^t = \beta^{-1} a \beta.$$  

(3.12.20)

If $k$ is a field, then the invertibility of $\beta$ in $M_n(k)$ is equivalent to the nondegeneracy of the corresponding bilinear form (3.12.1) on $k^n$ as a vector space over $k$, as in Section 2.14. In this case, (3.12.20) characterizes $T_b$ as the adjoint of $T_a$ with respect to (3.12.1).
3.13 Sesquilinear forms and matrices

Let \( n \) be a positive integer, so that the space \( \mathbb{C}^n \) of \( n \)-tuples of complex numbers is a vector space over \( \mathbb{C} \) with respect to coordinatewise addition and scalar multiplication, as usual. Also let \( (\beta_{j,l}) \) be an \( n \times n \) matrix with entries in \( \mathbb{C} \), and put

\[
(3.13.1) \quad \beta(z, w) = \sum_{j=1}^{n} \sum_{l=1}^{n} \beta_{j,l} z_l \bar{w}_j
\]

for every \( z, w \in \mathbb{C}^n \). This defines a sesquilinear form on \( \mathbb{C}^n \), and one can check that every sesquilinear form on \( \mathbb{C}^n \) corresponds to a unique matrix \( (\beta_{j,l}) \) in this way. Note that \( (3.13.1) \) is Hermitian-symmetric on \( \mathbb{C}^n \) if and only if

\[
(3.13.2) \quad \beta_{l,j} = \bar{\beta}_{j,l}
\]

for every \( j, l = 1, \ldots, n \).

If \( a = (a_{j,l}) \in M_n(\mathbb{C}) \) and \( z \in \mathbb{C}^n \), then let \( T_a(z) \) be the element of \( \mathbb{C}^n \) whose \( j \)th coordinate is given by

\[
(3.13.3) \quad (T_a(z))_j = \sum_{l=1}^{n} a_{j,l} z_l
\]

for each \( j = 1, \ldots, n \), as before. This defines a linear mapping from \( \mathbb{C}^n \) into itself, and \( a \mapsto T_a \) is an isomorphism from \( M_n(\mathbb{C}) \) as an algebra over \( \mathbb{C} \) with respect to matrix multiplication onto the algebra \( L(\mathbb{C}^n) \) of linear mappings from \( \mathbb{C}^n \) into itself with respect to composition of mappings, as in Section 2.10 and the previous section. Remember that \( a^* = ((a^*)_{j,l}) \in M_n(\mathbb{C}) \) is defined for each \( a \in M_n(\mathbb{C}) \) by

\[
(3.13.4) \quad (a^*)_{j,l} = \bar{a}_{l,j},
\]

as in Section 2.8, and that \( a \mapsto a^* \) is a conjugate-linear involution on \( M_n(\mathbb{C}) \).

Let \( a \in M_n(\mathbb{C}) \) be given, and observe that

\[
(3.13.5) \quad \beta(T_a(z), w) = \sum_{j=1}^{n} \sum_{l=1}^{n} \beta_{r,j} a_{j,l} z_l \bar{w}_r
\]

for every \( z, w \in \mathbb{C}^n \). Similarly,

\[
(3.13.6) \quad \beta(z, T_a(w)) = \sum_{j=1}^{n} \sum_{l=1}^{n} \beta_{l,r} a_{j,l} z_l \bar{w}_r
\]

for every \( z, w \in \mathbb{C}^n \). Thus \( T_a \) is self-adjoint with respect to \( \beta \), as in Section 2.15, if and only if

\[
(3.13.7) \quad \sum_{j=1}^{n} \beta_{r,j} a_{j,l} = \sum_{j=1}^{n} \beta_{l,j} \bar{a}_{j,r}
\]
for every $l, r = 1, \ldots, n$. If we also use $\beta$ to denote $(\beta_{j,l})$ as an element of $M_n(C)$, then (3.13.7) is the same as saying that

$$
(3.13.8) \quad \beta a = a^* \beta.
$$

Observe that

$$
(3.13.9) \quad \beta(T_a(z), T_a(w)) = \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{m=1}^{n} \sum_{r=1}^{n} \beta_{m,j} a_{j,l} \overline{a_{m,r}} z_l w_r
$$

for every $z, w \in C^n$. It follows that $T_a$ preserves $\beta(\cdot, \cdot)$, as in Section 3.11, if and only if

$$
(3.13.10) \quad \sum_{j=1}^{n} \sum_{m=1}^{n} \beta_{m,j} a_{j,l} \overline{a_{m,r}} = \delta_{r,l}
$$

for every $l, r = 1, \ldots, n$. This is the same as saying that

$$
(3.13.11) \quad a^* \beta a = \beta
$$

as elements of $M_n(C)$.

If $b \in M_n(C)$, then $T_b$ can be defined as in (3.13.3), and

$$
(3.13.12) \quad \beta(z, T_b(w)) = \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{r=1}^{n} \beta_{j,l} \overline{b_{j,r}} z_l w_r
$$

for every $z, w \in C^n$, as before. Hence

$$
(3.13.13) \quad \beta(T_a(z), w) = \beta(z, T_b(w))
$$

for every $z, w \in C^n$ if and only if

$$
(3.13.14) \quad \sum_{j=1}^{n} \beta_{r,j} a_{j,l} = \sum_{j=1}^{n} \beta_{j,l} \overline{b_{j,r}}
$$

for every $l, r = 1, \ldots, n$. This is the same as saying that

$$
(3.13.15) \quad \beta a = b^* \beta
$$

as elements of $M_n(C)$. Suppose that $(\beta_{j,l})$ is invertible as an element of $M_n(C)$, which is equivalent to the nondegeneracy of $\beta(\cdot, \cdot)$ as a sesquilinear form on $C^n$, as in Section 2.15. In this situation, (3.13.15) is the same as saying that

$$
(3.13.16) \quad b^* = \beta a \beta^{-1},
$$

which characterizes $T_b$ as the adjoint of $T_a$ with respect to $\beta(\cdot, \cdot)$. 


3.14 Invertibility and involutions

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an associative algebra over $k$ with a multiplicative identity element $e$. Also let $x \mapsto x^*$ be an opposite algebra automorphism on $A$, as in Section 2.6. Thus $e^* = e$, as before. If $x$ is an invertible element of $A$, then

\[(3.14.1)\quad x^* (x^{-1})^* = (x^{-1} x)^* = e^* = e \]

and

\[(3.14.2)\quad (x^{-1})^* x^* = (x x^{-1})^* = e^* = e. \]

This implies that $x^*$ is invertible in $A$, with

\[(3.14.3)\quad (x^*)^{-1} = (x^{-1})^*. \]

Let $\beta$ be an element of $A$, and let us say that $x \in A$ is *self-adjoint* with respect to $\beta$ and the given opposite algebra automorphism on $A$ if

\[(3.14.4)\quad \beta x = x^* \beta. \]

Similarly, let us say that $x$ is *anti-self-adjoint* with respect to $\beta$ and the given opposite algebra automorphism on $A$ if

\[(3.14.5)\quad \beta x = -x^* \beta. \]

If $\beta = e$, then these reduce to the usual notions of self-adjointness and anti-self-adjointness with respect to the given opposite algebra automorphism on $A$, as in Section 2.6. The collections of self-adjoint and anti-self-adjoint elements of $A$ with respect to $\beta$ and the given opposite algebra automorphism on $A$ are submodules of $A$, as a module over $k$.

Suppose for the moment that $x, y \in A$ are both anti-self-adjoint with respect to $\beta$ and the given opposite algebra automorphism on $A$. This implies that

\[(3.14.6)\quad \beta x y = -x^* \beta y = x^* y^* \beta = (y x)^* \beta \]

and

\[(3.14.7)\quad \beta y x = -y^* \beta x = y^* x^* \beta = (x y)^* \beta. \]

If $[x, y] = x y - y x$ is the usual commutator of $x$ and $y$ in $A$, then we get that

\[(3.14.8)\quad \beta [x, y] = \beta x y - \beta y x = (y x)^* \beta - (x y)^* \beta = -([x, y])^* \beta. \]

Thus $[x, y]$ is anti-self-adjoint with respect to $\beta$ and the given opposite algebra automorphism on $A$ too.

Suppose that $x, y \in A$ satisfy

\[(3.14.9)\quad x^* \beta x = \beta \]

and

\[(3.14.10)\quad y^* \beta y = \beta. \]
This implies that

\[(xy)^* \beta xy = y^* x^* \beta xy = y^* \beta y = \beta.\]

Note that (3.14.9) holds when \(x = e\), because \(e^* = e\). If \(x\) is an invertible element of \(A\) that satisfies (3.14.9), then

\[(3.14.12) \quad \beta = (x^*)^{-1} \beta x^{-1} = (x^{-1})^* \beta x^{-1},\]

using (3.14.3) in the second step. This shows that the collection of invertible elements \(x\) of \(A\) that satisfy (3.14.9) forms a group with respect to multiplication.

If \(\beta\) is an invertible element of \(A\), then

\[(3.14.13) \quad \phi_\beta(x) = \beta^{-1} x^* \beta\]
defines an opposite algebra automorphism on \(A\). In this case, (3.14.4) is equivalent to

\[(3.14.14) \quad \phi_\beta(x) = x,\]

which means that \(x\) is self-adjoint with respect to \(\phi_\beta\). Similarly, (3.14.5) is equivalent to

\[(3.14.15) \quad \phi_\beta(x) = -x,\]

which means that \(x\) is anti-self-adjoint with respect to \(\phi_\beta\). We also have that (3.14.9) is equivalent to

\[(3.14.16) \quad \phi_\beta(x) x = e,\]

which is the same as saying that

\[(3.14.17) \quad x^{-1} = \phi_\beta(x)\]

when \(x\) is invertible in \(A\). Of course, (3.14.13) reduces to the given opposite algebra automorphism \(x \mapsto x^*\) on \(A\) when \(\beta = e\).

Suppose for the moment that \(x \mapsto x^*\) is an involution on \(A\), as in Section 2.6. If \(x, \beta \in A\), then

\[(3.14.18) \quad (\beta x)^* = x^* \beta^*\]

and

\[(3.14.19) \quad (x^* \beta)^* = \beta^* (x^*)^* = \beta^* x.\]

It follows that \(x\) is self-adjoint or anti-self-adjoint with respect to \(\beta\) and the given involution on \(A\) if and only if \(x\) is self-adjoint or anti-self-adjoint, respectively, with respect to \(\beta^*\) and the given involution on \(A\). We also have that

\[(3.14.20) \quad (x^* \beta x)^* = x^* \beta^* (x^*)^* = x^* \beta x\]

for every \(x, \beta \in A\), so that (3.14.9) is equivalent to

\[(3.14.21) \quad x^* \beta^* x = \beta^* .\]
Let $\beta$ be an invertible element of $A$ again, so that $\phi_\beta$ can be defined on $A$ as in (3.14.13). Observe that
\[(\phi_\beta(x))^* = (\beta^{-1} x^* \beta)^* = \beta^* (x^*)^* (\beta^{-1})^* = \beta^* x (\beta^*)^{-1}\]
for every $x \in A$, using the hypothesis that $x \mapsto x^*$ be an involution on $A$ and (3.14.3) for $\beta$ in the last step. Thus
\[(\phi_\beta(\phi_\beta(x))) = \beta^{-1} \beta^* x (\beta^*)^{-1} \beta\]
for every $x \in A$. If $\beta$ is either self-adjoint or anti-self-adjoint with respect to the given involution on $A$, then it follows that $\phi_\beta$ is an involution on $A$ as well.

Suppose now that $k$ is the field $\mathbb{C}$ of complex numbers, and that $x \mapsto x^*$ is a conjugate-linear opposite algebra automorphism on $A$, as in Section 2.6. Thus $x \mapsto x^*$ may be considered as a real-linear opposite algebra automorphism on $A$ as an algebra over $\mathbb{R}$, as before. In particular, if $\beta \in A$, then the collections of elements of $A$ that are self-adjoint or anti-self-adjoint with respect to $\beta$ and $x \mapsto x^*$ are real-linear subspaces of $A$. In this situation, $x \in A$ is anti-self-adjoint with respect to $\beta$ if and only if $ix$ is self-adjoint with respect to $\beta$.

## 3.15 Invertibility and seminorms

Let $k$ be a field with an absolute value function $|\cdot|$, and let $A$ be an associative algebra over $k$ with a submultiplicative seminorm $N_A$ with respect to $|\cdot|$ on $k$. Suppose that $A$ has a multiplicative identity element $e$ such that
\[(3.15.1) \quad N_A(e) = 1.\]
If $x$ is an invertible element of $A$, then it follows that
\[(3.15.2) \quad N_A(x) N_A(x^{-1}) \geq 1.\]
In particular, if we have that
\[(3.15.3) \quad N_A(x), N_A(x^{-1}) \leq 1,\]
then
\[(3.15.4) \quad N_A(x) = N_A(x^{-1}) = 1.\]
Let $y$ be another invertible element of $A$ such that
\[(3.15.5) \quad N_A(y), N_A(y^{-1}) \leq 1.\]
Thus $xy$ is an invertible element of $A$ too,
\[(3.15.6) \quad N_A(x,y) \leq N_A(x) N_A(y) \leq 1,\]
and
\[(3.15.7) \quad N_A((xy)^{-1}) = N_A(y^{-1} x^{-1}) \leq N_A(x^{-1}) N_A(y^{-1}) \leq 1.\]
This shows that the collection of invertible elements $x$ of $A$ that satisfy (3.15.3) forms a group with respect to multiplication.

Suppose that $N_A$ is a semi-ultranorm on $A$, and that $x$ is an invertible element of $A$ that satisfies (3.15.3). Let $y$ be another invertible element of $A$ such that

$$N_A(x - y) < 1.$$  

(3.15.8)

This implies that

$$N_A(y) \leq \max(N_A(x), N_A(x - y)) \leq 1.$$  

(3.15.9)

We also have that

$$N_A(y^{-1}) = N_A(x^{-1}) \leq 1,$$  

(3.15.10)

as in (3.4.9). Thus $y$ satisfies (3.15.5) under these conditions.

Let $A$ be an algebra over $k$ in the strict sense, and let $N_A$ be a seminorm on $A$ with respect to $|\cdot|$ on $k$. Also let $x \mapsto x^*$ be an opposite algebra automorphism on $A$, as in Section 2.6. A basic compatibility condition between $x \mapsto x^*$ and $N_A$ is that there be a nonnegative real number $C_1$ such that

$$N_A(x^*) \leq C_1 N_A(x)$$  

(3.15.11)

for every $x \in A$. This is the same as saying that $x \mapsto x^*$ is bounded as a linear mapping from $A$ into itself, using $N_A$ on the domain and range. Another compatibility condition is that there be a nonnegative real number $C_2$ such that

$$N_A(x) \leq C_2 N_A(x^*)$$  

(3.15.12)

for every $x \in A$. In particular, (3.15.11) and (3.15.12) hold with $C_1 = C_2 = 1$ if and only if

$$N_A(x^*) = N_A(x)$$  

(3.15.13)

for every $x \in A$, which is to say that $x \mapsto x^*$ is an isometric linear mapping from $A$ into itself with respect to $N_A$. If $x \mapsto x^*$ is an involution on $A$, then (3.15.11) and (3.15.12) are equivalent, and with the same constant. If $k$ is the field of complex numbers with the standard absolute value function, then one can consider conjugate-linear opposite algebra automorphisms on $A$ as well.
Chapter 4

Formal power series

4.1 Direct sums and products

Let $k$ be a commutative ring with a multiplicative identity element, and let $I$ be a nonempty set. Suppose that for each $j \in I$, $A_j$ is a module over $k$. Under these conditions, the Cartesian product $\prod_{j \in I} A_j$ of the $A_j$’s is a module over $k$ too, with respect to coordinatewise addition and scalar multiplication. This is the direct product of the $A_j$’s, $j \in I$. If $a \in \prod_{j \in I} A_j$ and $l \in I$, then we let $a_l$ be the $l$th coordinate of $a$ in $A_l$. Thus $a \mapsto a_l$ is the standard coordinate projection from $\prod_{j \in I} A_j$ onto $A_l$. Of course, this mapping is linear over $k$.

Let $\bigoplus_{j \in I} A_j$ be the set of $a \in \prod_{j \in I} A_j$ such that $a_l = 0$ for all but finitely many $l \in I$. This is the direct sum of $A_j$, $j \in I$. Note that $\bigoplus_{j \in I} A_j$ is a submodule of $\prod_{j \in I} A_j$, as a module over $k$. If $I$ has only finitely many elements, then $\bigoplus_{j \in I} A_j$ is the same as $\prod_{j \in I} A_j$. If $I = \{1, \ldots, n\}$ for some $n \in \mathbb{Z}_+$, then we may use the notation $\bigoplus_{j=1}^n A_j$ or $\prod_{j=1}^n A_j$.

If $A_j$ is an algebra over $k$ in the strict sense for each $j \in I$, then $\prod_{j \in I} A_j$ is an algebra in the strict sense over $k$ with respect to coordinatewise multiplication, and $\bigoplus_{j \in I} A_j$ is a two-sided ideal in $\prod_{j \in I} A_j$. If $A_j$ is commutative for every $j \in I$, then $\prod_{j \in I} A_j$ is commutative as well. Similarly, if $A_j$ is associative for every $j \in I$, then $\prod_{j \in I} A_j$ is associative. If $A_j$ has a multiplicative identity element for every $j \in I$, then we get a multiplicative identity element in $\prod_{j \in I} A_j$. If $A_j$ is a Lie algebra for every $j \in I$, then $\prod_{j \in I} A_j$ is a Lie algebra.

Now let $k$ be a field with an absolute value function $|\cdot|$, and let $I$ be a nonempty set again. Suppose that $V_j$ is a vector space over $k$ for each $j \in I$, so that the direct product $\prod_{j \in I} V_j$ is a vector space over $k$ too. If $N_l$ is a seminorm on $V_l$ with respect to $|\cdot|$ on $k$ for some $l \in I$, then it is easy to see that

\begin{equation}
N_l(v) = N_l(v_l)
\end{equation}

defines a seminorm on $\prod_{j \in I} V_j$ with respect to $|\cdot|$ on $k$. If $N_l$ is a semi-ultranorm on $V_l$, then (4.1.1) is a semi-ultranorm on $\prod_{j \in I} V_j$. 

77
CHAPTER 4. FORMAL POWER SERIES

As before, the direct sum $\bigoplus_{j \in I} V_j$ is a linear subspace of $\prod_{j \in I} V_j$. Suppose that $N_j$ is a seminorm on $V_j$ for each $j \in I$, and put

\[(4.1.2) \quad \tilde{N}^1(v) = \sum_{j \in I} N_j(v_j)\]

for every $v \in \bigoplus_{j \in I} V_j$. More precisely, if $v \in \bigoplus_{j \in I} V_j$, then $v_j = 0$ for all but finitely many $j \in I$, so that the sum on the right side of (4.1.2) reduces to a finite sum of nonnegative real numbers. One can check that (4.1.2) defines a seminorm on $\bigoplus_{j \in I} V_j$ with respect to $| \cdot |$ on $k$, which is a norm when $N_j$ is a norm on $V_j$ for every $j \in I$.

Similarly, put

\[(4.1.3) \quad \tilde{N}^\infty(v) = \max_{j \in I} N_j(v_j)\]

for every $v \in \bigoplus_{j \in I} V_j$, which reduces to the maximum of finitely many nonnegative real numbers. One can verify that this defines a seminorm on $\bigoplus_{j \in I} V_j$, which is a norm when $N_j$ is a norm on $V_j$ for every $j \in I$. If $N_j$ is a semi-ultranorm on $V_j$ for each $j \in I$, then (4.1.3) is a semi-ultranorm on $\bigoplus_{j \in I} V_j$.

Observe that

\[(4.1.4) \quad \tilde{N}^1(v) \leq \tilde{N}^\infty(v)\]

for every $v \in \bigoplus_{j \in I} V_j$. If $I$ has only finitely many elements, then

\[(4.1.5) \quad \tilde{N}^1(v) \leq (#I) \tilde{N}^\infty(v)\]

for every $v \in \bigoplus_{j \in I} V_j$, where $#I$ is the number of elements in $I$.

Let $a_j$ be a positive real number for each $j \in I$. As before,

\[(4.1.6) \quad \tilde{N}^1_a(v) = \sum_{j \in I} a_j N_j(v_j)\]

defines a seminorm on $\bigoplus_{j \in I} V_j$ with respect to $| \cdot |$ on $k$. Similarly,

\[(4.1.7) \quad \tilde{N}^\infty_a(v) = \max_{j \in I} (a_j N_j(v_j))\]

defines a seminorm on $\bigoplus_{j \in I} V_j$ with respect to $| \cdot |$ on $k$, which is a semi-ultranorm when $N_j$ is a semi-ultranorm on $V_j$ for each $j \in I$. Clearly

\[(4.1.8) \quad \tilde{N}^\infty_a(v) \leq \tilde{N}^1_a(v)\]

for every $v \in \bigoplus_{j \in I} V_j$. If $I$ has only finitely many elements, then

\[(4.1.9) \quad \tilde{N}^1_a(v) \leq \left( \sum_{j \in I} a_j \right) \tilde{N}^\infty(v)\]

for every $v \in \bigoplus_{j \in I} V_j$. If $I$ has infinitely many elements, then $\sum_{j \in I} a_j$ can be defined as an extended real number, as the supremum of the corresponding finite subsums. If the supremum is finite, then (4.1.9) still holds and is nontrivial.
4.2 BILINEAR MAPPINGS AND CAUCHY PRODUCTS

Let \( k \) be a commutative ring with a multiplicative identity element, let \( A, B, \) and \( C \) be modules over \( k \), and let \( \beta \) be a mapping from \( A \times B \) into \( C \) that is bilinear over \( k \). Also let \( \sum_{j=0}^\infty a_j \) and \( \sum_{l=0}^\infty b_l \) be infinite series with terms in \( A \) and \( B \), respectively, considered formally for the moment. Put

\[
(4.2.1) \quad c_n = \sum_{j=0}^n \beta(a_j, b_{n-j})
\]

for each nonnegative integer \( n \). It is easy to see that

\[
(4.2.2) \quad \sum_{n=0}^\infty c_n = \beta\left( \sum_{j=0}^\infty a_j, \sum_{l=0}^\infty b_l \right),
\]

at least formally. More precisely, suppose for the moment that there are non-negative integers \( J, L \) such that \( a_j = 0 \) when \( j > J \) and \( b_l = 0 \) when \( l > L \). If \( n > J + L \), then it follows that \( c_n = 0 \). Thus the infinite series \( \sum_{j=0}^\infty a_j, \sum_{l=0}^\infty b_l \), and \( \sum_{n=0}^\infty c_n \) reduce to the finite sums \( \sum_{j=0}^J a_j, \sum_{l=0}^L b_l, \) and \( \sum_{n=0}^{J+L} c_n \), respectively, and the formal argument for (4.2.2) works in this case.

In particular, if \( A \) is an algebra over \( k \) in the strict sense, then we can take \( \beta \) to be the corresponding mapping from \( A \times A \) into \( A \). Let \( \sum_{j=0}^\infty a_j \) and \( \sum_{l=0}^\infty b_l \) be infinite series with terms in \( A \), and let us express multiplication of \( a, b \in A \) as \( ab \). The Cauchy product of these series is defined to be the series \( \sum_{n=0}^\infty c_n \), where

\[
(4.2.3) \quad c_n = \sum_{j=0}^n a_j b_{n-j},
\]

as in (4.2.1). Thus

\[
(4.2.4) \quad \sum_{n=0}^\infty c_n = \left( \sum_{j=0}^\infty a_j \right) \left( \sum_{l=0}^\infty b_l \right),
\]

at least formally, as before.

Suppose for the moment that \( k = \mathbb{R} \) with the standard absolute value function, and that \( a_j, b_l \) are nonnegative real numbers for every \( j, l \geq 0 \). If \( c_n \) is as in (4.2.3), then \( c_n \) is a nonnegative real number for every \( n \geq 0 \). If \( J, L \) are nonnegative integers, then one can verify that

\[
(4.2.5) \quad \left( \sum_{j=0}^J a_j \right) \left( \sum_{l=0}^L b_l \right) \leq \sum_{n=0}^{J+L} c_n.
\]

Similarly, if \( N \) is a nonnegative integer, then

\[
(4.2.6) \quad \sum_{n=0}^N c_n \leq \left( \sum_{j=0}^N a_j \right) \left( \sum_{l=0}^N b_l \right).
\]
If $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ converge, then it follows that $\sum_{n=0}^{\infty} c_n$ converges, and that the sums satisfy (4.2.4).

Let $k$ be any field with an absolute value function $|\cdot|$, and let $A$, $B$, and $C$ be vector spaces over $k$ with norms $N_A$, $N_B$, and $N_C$, respectively, with respect to $|\cdot|$ on $k$. Also let $\beta$ be a bounded bilinear mapping from $A \times B$ into $C$ with respect to these norms, so that there is a nonnegative real number $C(\beta)$ such that

$$N_C(\beta(a, b)) \leq C(\beta) N_A(a) N_B(b)$$

for every $a \in A$ and $b \in B$. Suppose that $A$, $B$, and $C$ are complete with respect to the metrics associated to $N_A$, $N_B$, and $N_C$, respectively. Let $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ be infinite series with terms in $A$ and $B$ that converge absolutely with respect to $N_A$ and $N_B$, respectively, so that $\sum_{j=0}^{\infty} N_A(a_j)$ and $\sum_{l=0}^{\infty} N_B(b_l)$ converge as infinite series of nonnegative real numbers. If $c_n$ is as in (4.2.1), then

$$N_C(c_n) \leq \sum_{j=0}^{n} N_C(\beta(a_j, b_{n-j})) \leq C(\beta) \sum_{j=0}^{n} N_A(a_j) N_B(b_{n-j})$$

for every $n \geq 0$. The sum on the right side of (4.2.8) is the same as the $n$th term of the Cauchy product of $\sum_{j=0}^{\infty} N_A(a_j)$ and $\sum_{l=0}^{\infty} N_B(b_l)$. It follows that $\sum_{n=0}^{\infty} N_C(c_n)$ converges as an infinite series of nonnegative real numbers, with

$$\sum_{n=0}^{\infty} N_C(c_n) \leq C(\beta) \left( \sum_{j=0}^{\infty} N_A(a_j) \right) \left( \sum_{l=0}^{\infty} N_B(b_l) \right).$$

Thus $\sum_{n=0}^{\infty} c_n$ converges absolutely with respect to $N_C$, and $\sum_{j=0}^{\infty} a_j$, $\sum_{l=0}^{\infty} b_l$, and $\sum_{n=0}^{\infty} c_n$ converge in $A$, $B$, and $C$, respectively, by completeness. One can check that (4.2.2) holds in this situation, by approximating these sums by finite sums.

Suppose now that $N_A$, $N_B$, and $N_C$ are ultranorms on $A$, $B$, and $C$, respectively, and let $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ be infinite series with terms in $A$ and $B$, respectively, such that

$$\lim_{j \to \infty} N_A(a_j) = \lim_{l \to \infty} N_B(b_l) = 0.$$ 

If $c_n$ is as in (4.2.1) again, then

$$N_C(c_n) \leq \max_{0 \leq j \leq n} N_C(\beta(a_j, b_{n-j})) \leq C(\beta) \max_{0 \leq j \leq n} (N_A(a_j) N_B(b_l))$$

for every $n \geq 0$. One can verify that

$$\lim_{n \to \infty} N_C(c_n) = 0,$$

using (4.2.10) and (4.2.11). It follows that $\sum_{j=0}^{\infty} a_j$, $\sum_{l=0}^{\infty} b_l$, and $\sum_{n=0}^{\infty} c_n$ converge in $A$, $B$, and $C$, respectively, because of completeness, as in Section 1.12. One can check that (4.2.2) holds in this situation too, by approximating these sums by finite sums, as before.
4.3 Formal power series and modules

Let $k$ be a commutative ring with a multiplicative identity element, let $A$ be a module over $k$, and let $T$ be an indeterminate. As in [4, 10], we shall try to use upper-case letters like $T$ for indeterminates, and lower-case letters for elements of $k$ or $A$. A formal power series in $T$ with coefficients in $A$ can be expressed as

\[(4.3.1)\quad f(T) = \sum_{j=0}^{\infty} f_j T^j,\]

where $f_j$ is an element of $A$ for each nonnegative integer $j$. The space $A[[T]]$ of these formal power series can be defined as the space of functions on the set $\mathbb{Z}_+ \cup \{0\}$ of nonnegative integers with values in $A$, where (4.3.1) corresponds to $j \mapsto f_j$ as an $A$-valued function on $\mathbb{Z}_+ \cup \{0\}$. This is a module over $k$ with respect to pointwise addition and scalar multiplication of $A$-valued functions on $\mathbb{Z}_+ \cup \{0\}$, which corresponds to termwise addition and scalar multiplication of formal power series as in (4.3.1).

Similarly, a formal polynomial in $T$ with coefficients in $A$ can be expressed as

\[(4.3.2)\quad f(T) = \sum_{j=0}^{n} f_j T^j\]

for some nonnegative integer $n$, where $f_j \in A$ for each $j = 0, \ldots, n$. This may be considered as a formal power series in $T$ too, with $f_j = 0$ for $j > n$. The space $A[T]$ of these formal polynomials can be defined as the space of $A$-valued functions on $\mathbb{Z}_+ \cup \{0\}$ that are equal to 0 at all but finitely many nonnegative integers. Of course, $A[T]$ is a submodule of $A[[T]]$, as a module over $k$. Observe that $A[[T]]$ corresponds to the direct product of copies of $A$ indexed by the set $\mathbb{Z}_+ \cup \{0\}$ of nonnegative integers, as a module over $k$, and that $A[T]$ corresponds to the analogous direct sum.

There is a natural mapping from $A$ into $A[T]$, which sends $a \in A$ to the formal polynomial $f(T)$ with $f_0 = a$ and $f_j = 0$ when $j \geq 1$. This is an injective module homomorphism from $A$ into $A[T]$, and it is sometimes convenient to think of $A$ as a submodule of $A[T]$ in this way. Note that the mapping

\[(4.3.3)\quad f(T) \mapsto f_0\]

is a module homomorphism from $A[[T]]$ onto $A$.

If $f(T) \in A[[T]]$ and $l$ is a nonnegative integer, then

\[(4.3.4)\quad f(T) T^l = \sum_{j=0}^{\infty} f_j T^{j+l} = \sum_{j=l}^{\infty} f_{j-l} T^j\]

defines an element of $A[[T]]$, which is the same as $f(T)$ when $l = 0$. Of course, (4.3.4) is in $A[T]$ when $f(T) \in A[T]$. The mapping

\[(4.3.5)\quad f(T) \mapsto f(T) T^l\]
is an injective module homomorphism from \( A[[T]] \) into itself for every \( l \geq 0 \).

Let us say that \( f(T) \in A[[T]] \) vanishes to order \( n \) for some nonnegative integer \( n \) if \( f_j = 0 \) for every \( j = 0, \ldots, n \). Equivalently, this means that \( f(T) = g(T) T^{n+1} \) for some \( g(T) \in A[[T]] \). The collection of \( f(T) \in A[[T]] \) that vanishes to order \( n \) is a submodule of \( A[[T]] \). This submodule is the same as the kernel of (4.3.3) when \( n = 0 \).

### 4.4 Sums and extensions

Let \( k \) be a commutative ring with a multiplicative identity element, let \( A \) be a module over \( k \), and let \( T \) be an indeterminate again. Also let \( f_l(T) = \sum_{j=0}^{\infty} f_{l,j} T^j \) be an element of \( A[[T]] \) for each \( l \in \mathbb{Z}_+ \), and let \( f(T) \) be another element of \( A[[T]] \). Let us say that \( \{f_l(T)\}_{l=1}^{\infty} \) eventually agrees termwise with \( f(T) \) termwise if for each nonnegative integer \( j \) there is a positive integer \( L_j \) such that

\[
(4.4.1) \quad f_{l,j} = f_j
\]

for every \( l \geq L_j \). This implies that \( f_l(T) - f(T) \) vanishes to order \( n \) when \( l \geq \max(L_0, \ldots, L_n) \). As in the previous section, \( A[[T]] \) can be defined as the space of \( A \)-valued functions on \( \mathbb{Z}_+ \cup \{0\} \), which is the same as the Cartesian product of the family of copies of \( A \) indexed by \( \mathbb{Z}_+ \cup \{0\} \). Consider the product topology on \( A[[T]] \) as a Cartesian product, using the discrete topology on \( A \) in each factor. The condition that \( \{f_l(T)\}_{l=1}^{\infty} \) eventually agree termwise with \( f(T) \) is equivalent to the convergence of \( \{f_l(T)\}_{l=1}^{\infty} \) to \( f(T) \) with respect to this product topology.

Similarly, let us say that \( \{f_l(T)\}_{l=1}^{\infty} \) is termwise eventually constant if for each nonnegative integer \( j \) there is an \( L_j \in \mathbb{Z}_+ \) such that \( f_{l,j} \) does not depend on \( l \) when \( l \geq L_j \). In this case, we can define \( f_j \in A \) for each \( j \geq 0 \) by putting \( f_j = f_{l,j} \) when \( l \geq L_j \). This defines \( f(T) = \sum_{j=0}^{\infty} f_j T^j \) as an element of \( A[[T]] \), and we have that \( \{f_l(T)\}_{l=1}^{\infty} \) eventually agrees termwise with \( f(T) \). Conversely, if \( \{f_l(T)\}_{l=1}^{\infty} \) eventually agrees termwise with some \( f(T) \in A[[T]] \), then \( \{f_l(T)\}_{l=1}^{\infty} \) is termwise eventually constant.

If \( \{f_l(T)\}_{l=1}^{\infty} \) eventually agrees termwise with some \( f(T) \in A[[T]] \) and \( \alpha \in k \), then \( \{\alpha f_l(T)\}_{l=1}^{\infty} \) eventually agrees termwise with \( \alpha f(T) \). In this situation, we also have that \( \{f_l(T) T^r\}_{l=1}^{\infty} \) eventually agrees termwise with \( f(T) T^r \) for every nonnegative integer \( r \). If \( \{g_l(T)\}_{l=1}^{\infty} \) is another sequence of elements of \( A[[T]] \) that eventually agrees termwise with \( g(T) \in A[[T]] \), then \( \{f_l(T) + g_l(T)\}_{l=1}^{\infty} \) eventually agrees termwise with \( f(T) + g(T) \).

Let \( a_l(T) = \sum_{j=0}^{\infty} a_{l,j} T^j \) be a formal power series in \( T \) with coefficients in \( A \) for each \( l \in \mathbb{Z}_+ \). Suppose that \( \{a_l(T)\}_{l=1}^{\infty} \) eventually agrees termwise with 0, so that for each \( j \geq 0 \) there is an \( L_j \in \mathbb{Z}_+ \) such that \( a_{l,j} = 0 \) for every \( l \geq L_j \). It follows that the coefficient of \( T^j \) in

\[
\sum_{l=1}^{n} a_l(T)
\]
does not depend on \( n \) when \( n \geq L_j - 1 \), so that the sequence of these sums is termwise eventually constant. Under these conditions, we can define

\[
\sum_{l=1}^{\infty} a_l(T)
\]

as a formal power series in \( T \) with coefficients in \( A \), by taking the coefficient of \( T^j \) in (4.4.3) to be the coefficient of \( T^j \) in (4.4.2) when \( l \geq L(n) - 1 \), as before. By construction, the sequence of partial sums (4.4.2) eventually agrees termwise with (4.4.3).

If \( \alpha \in k \), then \( \{ \alpha a_l(T) \}_{i=1}^{\infty} \) eventually agrees termwise with 0 too, and

\[
\sum_{l=1}^{\infty} \alpha a_l(T) = \alpha \sum_{l=1}^{\infty} a_l(T).
\]

If \( r \) is a nonnegative integer, then \( \{ a_l(T) T^r \}_{i=1}^{\infty} \) eventually agrees termwise with 0 as well, and

\[
\sum_{l=1}^{\infty} a_l(T) T^r = \left( \sum_{l=1}^{\infty} a_l(T) \right) T^r.
\]

Let \( B \) be another module over \( k \), so that \( B[[T]] \) is a module over \( k \) too, as before. Also let \( \phi \) be a module homomorphism from \( A \) into \( B[[T]] \). Note that a module homomorphism from \( A \) into \( B \) may be considered as a module homomorphism from \( A \) into \( B[[T]] \), by considering \( B \) as a submodule of \( B[[T]] \). If \( f(T) = \sum_{i=0}^{\infty} f_i T^i \in A[[T]] \), then \( \phi(f(T)) = \sum_{i=0}^{\infty} \phi(f_i) T^i \) for each \( l \geq 0 \), and \( \phi(f(T)) T^l \) automatically vanishes to order \( l - 1 \) for every \( l \geq 1 \). In particular, \( \{ \phi(f(T)) T^l \}_{l=1}^{\infty} \) eventually agrees termwise with 0, and we put

\[
\phi(f(T)) = \sum_{l=0}^{\infty} \phi(f_l) T^l,
\]

where the sum is defined as an element of \( B[[T]] \) as in (4.4.3). This defines a module homomorphism from \( A[[T]] \) into \( B[[T]] \), which agrees with the initial homomorphism from \( A \) into \( B[[T]] \) when \( A \) is considered as a submodule of \( A[[T]] \). If we start with a module homomorphism \( \phi \) from \( A \) into \( B[[T]] \), and if \( f(T) \in A[T] \), then (4.4.7) reduces to a finite sum in \( B[T] \).

If \( f(T) \in A[[T]] \) vanishes to order \( n \) for some nonnegative integer \( n \), then it is easy to see that \( \phi(f(T)) \) vanishes to order \( n \) as well. More precisely, one can check that

\[
\phi(f(T) T^r) = \phi(f(T)) T^r
\]
for every \( f(T) \in A[[T]] \) and nonnegative integer \( r \). If \( f(T) = \sum_{j=0}^{\infty} f_j T^j \in A[[T]] \) and \( j \) is a nonnegative integer, then the total coefficient of \( T^j \) in (4.4.7) is the sum of the coefficients of \( T^j \) in \( \phi(f_l) \) for \( l = 0, \ldots, j \), and in particular depends only on \( \phi(f_l) \) for \( l \leq j \). If \( \{f_r(T)\}_{r=1}^{\infty} \) is a sequence of elements of \( A[[T]] \) that eventually agrees termwise with \( f(T) \), then it follows that \( \{\phi(f_r(T))\}_{r=1}^{\infty} \) eventually agrees termwise with \( \phi(f(T)) \).

### 4.5 Extending bilinear mappings

Let \( k \) be a commutative ring with a multiplicative identity element, let \( A, B, \) and \( C \) be modules over \( k \), and let \( T \) be an indeterminate. As before, \( C[[T]] \) is a module over \( k \), and we let \( \beta \) be a mapping from \( A \times B \) into \( C[[T]] \) that is bilinear over \( k \). There is a natural way to extend \( \beta \) to a mapping from \( A[[T]] \times B[[T]] \) into \( C[[T]] \), as follows. Let \( f(T) = \sum_{j=0}^{\infty} f_j T^j \) and \( g(T) = \sum_{l=0}^{\infty} g_l T^l \) be formal power series in \( T \) with coefficients in \( A \) and \( B \), respectively. Put

\[
(4.5.1) \quad h_n(T) = \sum_{j=0}^{n} \beta(f_j, g_{n-j})
\]

for each nonnegative integer \( n \), which is an element of \( C[[T]] \). Thus \( h_n(T) T^n \) automatically vanishes to order \( n - 1 \) for every \( n \geq 1 \), so that

\[
(4.5.2) \quad h(T) = \sum_{n=0}^{\infty} h_n(T) T^n
\]

defines an element of \( C[[T]] \) too, as in the previous section. Put

\[
(4.5.3) \quad \beta(f(T), g(T)) = h(T).
\]

This defines a mapping from \( A[[T]] \times B[[T]] \) into \( C[[T]] \) that is bilinear over \( k \) and agrees with the initial mapping from \( A \times B \) into \( C[[T]] \), with \( A \) and \( B \) considered as submodules of \( A[[T]] \) and \( B[[T]] \), respectively.

If \( f(T) \in A[[T]] \) and \( g(T) \in B[[T]] \), then (4.5.1) is equal to 0 for all but finitely many \( n \), so that (4.5.2) reduces to a finite sum. If the initial mapping \( \beta \) sends \( A \times B \) into \( C[[T]] \), then (4.5.1) is in \( C[[T]] \) for every \( n \geq 0 \). In this case, it follows that (4.5.2) is an element of \( C[[T]] \) when \( f(T) \in A[[T]] \) and \( g(T) \in B[[T]] \).

If \( f(T) \in A[[T]] \) and \( g(T) \in B[[T]] \) vanish to order \( r_1 \) and \( r_2 \), respectively, for some nonnegative integers \( r_1, r_2 \), then (4.5.1) is equal to 0 when \( n \leq r_1 + r_2 \), so that (4.5.2) vanishes to order \( r_1 + r_2 \). More precisely, one can verify that

\[
(4.5.4) \quad \beta(f(T) T^r, g(T) T^m) = \beta(f(T), g(T)) T^{r+m}
\]

for every \( f(T) \in A[[T]] \), \( g(T) \in B[[T]] \), and nonnegative integers \( r, m \). If \( f(T) = \sum_{j=0}^{\infty} f_j T^j \in A[[T]] \), \( g(T) = \sum_{l=0}^{\infty} g_l T^l \in B[[T]] \), and \( r \) is a nonnegative integer, then the total coefficient of \( T^r \) in (4.5.2) is the sum of the coefficients of \( T^{r-n} \) in (4.5.1) for \( n = 0, \ldots, r \). In particular, this only involves (4.5.1) for \( n \leq r \), and
4.5. EXTENDING BILINEAR MAPPINGS

hence depends only on \( f_j \) and \( g_l \) for \( j, l \leq r \). If \( \{f_m(T)\}_{m=1}^{\infty} \) and \( \{g_m(T)\}_{m=1}^{\infty} \) are sequences of elements of \( A[[T]] \) and \( B[[T]] \), respectively, then it follows that \( \{\beta(f_m(T), g_m(T))\}_{m=1}^{\infty} \) eventually agrees termwise with \( \beta(f(T), g(T)) \).

Let \( \{a_m(T)\}_{m=0}^{\infty} \) and \( \{b_r(T)\}_{r=0}^{\infty} \) be sequences of elements of \( A[[T]] \) and \( B[[T]] \), respectively, that eventually agree termwise with \( f(T) \) and \( g(T) \), respectively, then it follows that \( \{\beta(a_m(T), b_r(T))\}_{m=0}^{\infty} \) eventually agrees termwise with \( \beta(f(T), g(T)) \).

(4.5.5)

\[
c_N(T) = \sum_{m=0}^{N} \beta(a_m(T), b_{N-m}(T))
\]

for every nonnegative integer \( N \), which is an element of \( C[[T]] \). Note that \( \beta(a_m(T), b_r(T)) \) vanishes to arbitrarily large order when \( m \) or \( r \) is sufficiently large in this situation, by the remarks in the preceding paragraph. This implies that \( \{c_N(T)\}_{N=0}^{\infty} \) eventually agrees termwise with 0, so that \( \sum_{N=0}^{\infty} c_N(T) \) defines an element of \( C[[T]] \), as in the previous section. One can check that

(4.5.6)

\[
\sum_{N=0}^{\infty} c_N(T) = \beta \left( \sum_{m=0}^{\infty} a_m(T), \sum_{r=0}^{\infty} b_r(T) \right)
\]

as elements of \( C[[T]] \), as in Section 4.2. More precisely, this means that for each nonnegative integer \( j \), the coefficients of \( T^j \) on both sides of (4.5.6) are the same. This reduces to an analogous statement for finite sums for each \( j \geq 0 \) in this situation.

Observe that (4.5.1) is the same as

(4.5.7)

\[
\sum_{l=0}^{n} \beta(f_{n-l}, g_l).
\]

Suppose now that \( A = B \). If the initial mapping \( \beta \) from \( A \times A \) into \( C[[T]] \) is symmetric or antisymmetric, then it is easy to see that the extension of \( \beta \) to \( A[[T]] \times A[[T]] \) has the same property, using the fact that (4.5.1) is the same as (4.5.7). Suppose that the initial mapping \( \beta \) on \( A \times A \) satisfies

(4.5.8)

\[
\beta(a, a) = 0
\]

for every \( a \in A \), and let us check that the extension of \( \beta \) to \( A[[T]] \times A[[T]] \) satisfies

(4.5.9)

\[
\beta(f(T), f(T)) = 0
\]

for every \( f(T) = \sum_{j=0}^{\infty} f_j T^j \in A[[T]] \). To do this, it suffices to verify that

(4.5.10)

\[
\sum_{j=0}^{n} \beta(f_j, f_{n-j}) = 0
\]
for each nonnegative integer $n$. Remember that (4.5.8) implies that $\beta$ is anti-

symmetric on $A \times A$, as in Section 2.1. If $n$ is odd, then (4.5.10) reduces to the

antisymmetry of $\beta$ on $A \times A$. If $n$ is even, then one can use the antisymmetry

of $\beta$ on $A \times A$ to reduce (4.5.10) to the condition that $\beta(f_n/2, f_n/2) = 0$, which

follows from (4.5.8).

### 4.6 Formal power series and algebras

Let $k$ be a commutative ring with a multiplicative identity element again, and let

$T$ be an indeterminate. Suppose that $A$ is an algebra over $k$ in the strict sense,

where multiplication of $a, b \in A$ is expressed as $ab$. Let $f(T) = \sum_{j=0}^{\infty} f_j T^j$ and

g(T) = \sum_{t=0}^{\infty} g_t T^t$ be formal power series in $T$ with coefficients in $A$, and put

$$
(4.6.1) \quad h_n = \sum_{j=0}^{n} f_j g_{n-j}
$$

for each nonnegative integer $n$. Thus $h(T) = \sum_{n=0}^{\infty} h_n T^n$ is a formal power

series in $T$ with coefficients in $A$ too, and we put

$$
(4.6.2) \quad f(T) g(T) = h(T).
$$

This defines a mapping from $A[[T]] \times A[[T]]$ into $A[[T]]$, which is the same

as the mapping obtained from multiplication on $A$ as in the previous section.

Using this definition of multiplication on $A[[T]]$, we get that $A[[T]]$ is an algebra

over $k$ in the strict sense. Note that $A[[T]]$ is a subalgebra of $A[[T]]$, and that

$A$ corresponds to a subalgebra of $A[T]$, using the identification mentioned in

Section 4.3. The mapping $f(T) \mapsto f_0$ mentioned in Section 4.3 defines an algebra

homomorphism from $A[[T]]$ onto $A$. If multiplication on $A$ is commutative,

then multiplication on $A[[T]]$ is commutative as well, as in the remark about

symmetry of $\beta$ in the previous section. Similarly, if multiplication on $A$ is

associative, then one can check that multiplication on $A[[T]]$ is associative too.

If $A$ has a multiplicative identity element $e$, then the corresponding formal

polynomial in $T$ is the multiplicative identity element in $A[[T]]$.

In particular, we can take $A = k$, considered as an algebra over itself. Thus

$k[[T]]$ is a commutative associative algebra over $k$ with a multiplicative identity

element, and $k[T]$ is a subalgebra of $k[[T]]$. We can identify $k$ with a subalge-

bra of $k[T]$, where the multiplicative identity element in $k$ corresponds to the

multiplicative identity element in $k[[T]]$.

If $A$ is any module over $k$, then $A[[T]]$ may be considered as a module over

$k[[T]]$. More precisely, if $f(T) \in k[[T]]$ and $g(T) \in A[[T]]$, then $f(T) g(T)$ can

be defined as a formal power series in $T$ with coefficients in $A$ as in (4.6.2),

where the terms in the sum on the right side of (4.6.1) are defined using scalar

multiplication on $A$. Equivalently, scalar multiplication on $A$ corresponds to a

mapping from $k \times A$ into $A$ that is bilinear over $k$, which can be extended to a

mapping from $k[[T]] \times A[[T]]$ into $A[[T]]$ as in the previous section. Similarly,

$A[T]$ may be considered as a module over $k[T]$.
4.7. INVERTIBILITY IN $A[[T]]$

Let $A$ and $B$ be modules over $k$, and let $\phi$ be a module homomorphism from $A$ into $B[[T]]$, as modules over $k$. As in Section 4.4, there is a natural way to extend $\phi$ to a mapping from $A[[T]]$ into $B[[T]]$. It is easy to see that this mapping is a module homomorphism from $A[[T]]$ into $B[[T]]$, as modules over $k[[T]]$. If the initial mapping sends $A$ into $B[T]$, then the restriction to $A[T]$ of the extension to $A[[T]]$ is a module homomorphism into $B[T]$, as modules over $k[T]$. Let $C$ be another module over $k$, and let $\beta$ be a mapping from $A \times B$ into $C[[T]]$ that is bilinear over $k$. One can check that the extension of $\beta$ to a mapping from $A[[T]] \times B[[T]]$ into $C[[T]]$ defined in the previous section is bilinear over $k[[T]]$. If the initial mapping sends $A \times B$ into $C[T]$, then the restriction to $A[T] \times B[T]$ of the extension to $A[[T]] \times B[[T]]$ is bilinear over $k[T]$ as a mapping into $C[T]$. If $A$ is an algebra over $k$ in the strict sense, then $A[[T]]$ may be considered as an algebra over $k[[T]]$ in the strict sense, and $A[T]$ may be considered as an algebra over $k[T]$ in the strict sense.

Suppose that $(A, \langle \cdot, \cdot \rangle_A)$ is a Lie algebra over $k$, and let $f(T)$ and $g(T)$ be formal power series in $T$ with coefficients in $A$ again. In this situation, (4.6.1) should be expressed as

$$h_n = \sum_{j=0}^{n} [f_j, g_{n-j}]_A$$

for each $n \geq 0$, and we put

$$[f(T), g(T)]_{A[[T]]} = h(T) = \sum_{n=0}^{\infty} h_n T^n.$$

One can verify that $A[[T]]$ is a Lie algebra over $k$ with respect to (4.6.4). More precisely, one can use the fact that $[a, a]_A = 0$ for every $a \in A$ to get that $[f(T), f(T)]_{A[[T]]} = 0$ for every $f(T) \in A[[T]]$, as in the previous section. One can also get the Jacobi identity for $\langle \cdot, \cdot \rangle_{A[[T]]}$ from the Jacobi identity for $\langle \cdot, \cdot \rangle_A$. As before, $A[T]$ is a Lie subalgebra of $A[[T]]$, as a Lie algebra over $k$. One can consider $A[[T]]$ as a Lie algebra over $k[[T]]$, and $A[T]$ as a Lie algebra over $k[T]$.

4.7 Invertibility in $A[[T]]$

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an associative algebra over $k$ with a multiplicative identity element $e$. Also let $T$ be an indeterminate, so that $A[[T]]$ may be considered as an associative algebra over $k[[T]]$ as in the previous section. Let us identify $e$ with the corresponding formal power series in $T$ with coefficients in $A$, which is the multiplicative identity element in $A[[T]]$. If $a(T) \in A[[T]]$, then $a(T)^l$ can be defined as an element of $A[[T]]$ for every $l \in \mathbb{Z}_+$ using multiplication on $A[[T]]$, and we interpret $a(T)^l$ as being equal to $e$ when $l = 0$. Observe that

$$(4.7.1) \quad (e - a(T)) \sum_{l=0}^{n} a(T)^l = \left( \sum_{l=0}^{n} a(T)^l \right) (e - a(T)) = e - a(T)^{n+1}$$
for every nonnegative integer \( n \), by a standard computation.

Suppose that \( a(T) \) vanishes to order 0, as in Section 4.3, so that the coefficient of \( T^0 \) in \( a(T) \) is equal to 0. This implies that \((a(T))^l\) vanishes to order \( l - 1 \) for every \( l \in \mathbb{Z}_+ \). It follows that the coefficient of \( T^j \) in

\[
\sum_{l=0}^{n} a(T)^l
\]

is the same for \( n \geq j \). As in Section 4.4, we define

\[
\sum_{l=0}^{\infty} a(T)^l
\]

as a formal power series in \( T \) with coefficients in \( A \) by taking the coefficient of \( T^j \) in (4.7.3) to be the same as the coefficient of \( T^j \) in (4.7.2) when \( n \geq j \). In particular, \( \{a(T)^l\}_{l=0}^{\infty} \) eventually agrees termwise with 0, as in Section 4.4, and the sequence of partial sums (4.7.2) eventually agrees termwise with (4.7.3).

Using (4.7.1), one can check that

\[
(e - a(T)) \sum_{l=0}^{\infty} a(T)^l = (\sum_{l=0}^{\infty} a(T)^l)(e - a(T)) = e.
\]

More precisely, for each nonnegative integer \( j \), the coefficient of \( T^j \) in each of the three expressions in (4.7.4) is the same as in the corresponding expression in (4.7.1) when \( n \geq j \). It follows that (4.7.3) is the multiplicative inverse of \( e - a(T) \) in \( A[[T]] \).

Let \( f(T) = \sum_{j=0}^{\infty} a_j T^j \) be a formal power series in \( T \) with coefficients in \( A \). If \( f_0 \) is invertible as an element of \( A \), then \( f(T) \) can be expressed as \( f_0 (e - a(T)) \), where \( a(T) \in A[[T]] \) vanishes to order 0. This implies that \( f(T) \) is invertible in \( A[[T]] \), because \( e - a(T) \) is invertible in \( A[[T]] \), as in the previous paragraph. Conversely, if \( f(T) \) is invertible in \( A[[T]] \), then \( f_0 \) is invertible in \( A \), because \( f(T) \mapsto f_0 \) is an algebra homomorphism from \( A[[T]] \) onto \( A \).

Of course, the collection of invertible elements of \( A[[T]] \) is a group with respect to multiplication of formal power series. The collection of \( f(T) \in A[[T]] \) with \( f_0 = e \) is a subgroup of this group.

## 4.8 Homomorphisms over \( k[T] \)

Let \( k \) be a commutative ring with a multiplicative identity element, let \( A, B \) be modules over \( k \), and let \( T \) be an indeterminate. Also let \( \phi \) be a homomorphism from \( A[T] \) into \( B[T] \), as modules over \( k \). Suppose that

\[
\phi(f(T)) = \phi(f(T)) T
\]

for every \( f(T) \in A[T] \). This implies that

\[
\phi(f(T) T^n) = \phi(f(T)) T^n
\]
for every \( f(T) \in A[T] \) and \( r \in \mathbb{Z}_+ \), by applying (4.8.1) repeatedly. Of course, (4.8.2) holds trivially when \( r = 0 \). It follows that \( \phi \) is a homomorphism from \( A[T] \) into \( B[T] \) as modules over \( k[T] \) under these conditions. Conversely, if \( \phi \) is a homomorphism from \( A[T] \) into \( B[T] \) as modules over \( k[T] \), then \( \phi \) is a homomorphism from \( A[T] \) into \( B[T] \) as modules over \( k[T] \) that satisfies (4.8.1).

Let \( \phi \) be a homomorphism from \( A[T] \) into \( B[T] \) as modules over \( k[T] \) again. It is easy to see that \( \phi \) is uniquely determined on \( A[T] \) by its restriction to \( A \), considered as a submodule of \( A[T] \) as a module over \( k \), as in Section 4.3. Remember that every homomorphism from \( A \) into \( B[T] \), as modules over \( k \), can be extended to a homomorphism from \( A[T] \) into \( B[T] \) as modules over \( k[T] \), as in Sections 4.4 and 4.6. This gives a natural isomorphism between \( \text{Hom}_k(A, B[T]) \) and \( \text{Hom}_{k[T]}(A[T], B[T]) \), as modules over \( k \).

In fact, \( \text{Hom}_k(A, B[T]) \) may be considered as a module over \( k[T] \). More precisely, \( \text{Hom}_k(A, B[T]) \) is a submodule of the space of all \( B[T] \)-valued functions on \( A \), as a module over \( k[T] \). The isomorphism between \( \text{Hom}_k(A, B[T]) \) and \( \text{Hom}_{k[T]}(A[T], B[T]) \) mentioned in the preceding paragraph is linear over \( k[T] \).

Now let \( \phi \) be a homomorphism from \( A[[T]] \) into \( B[[T]] \), as modules over \( k \), that satisfies (4.8.1) for every \( f(T) \in A[[T]] \). This implies that (4.8.2) holds for every \( f(T) \in A[[T]] \) and nonnegative integer \( r \), as before. If \( f(T) \in A[[T]] \) vanishes to order \( n \) for some nonnegative integer \( n \), then we get that \( \phi(f(T)) \) vanishes to order \( n \) as well, by expressing \( f(T) \) as an element of \( A[[T]] \) times \( T^{n+1} \). Using this, one can check that \( \phi \) is a homomorphism from \( A[[T]] \) into \( B[[T]] \), as modules over \( k[[T]] \). Conversely, if \( \phi \) is a homomorphism from \( A[[T]] \) into \( B[[T]] \) as modules over \( k[[T]] \), then \( \phi \) is a homomorphism from \( A[[T]] \) into \( B[[T]] \) as modules over \( k[T] \), and hence \( \phi \) is a homomorphism from \( A[[T]] \) into \( B[[T]] \) as modules over \( k[T] \) that satisfies (4.8.1) for every \( f(T) \in A[[T]] \).

Let \( \phi \) be a homomorphism from \( A[[T]] \) into \( B[[T]] \) as modules over \( k \) that satisfies (4.8.1) for every \( f(T) \in A[[T]] \) again. If \( f(T) = \sum_{l=0}^{\infty} f_l T^l \in A[[T]] \) and \( j \) is a nonnegative integer, then

\[
\phi\left( \sum_{l=j+1}^{\infty} f_l T^l \right)
\]

vanishes to order \( j \), as before. This implies that the total coefficient of \( T^j \) in \( \phi(f(T)) \) is the same as for

\[
\phi\left( \sum_{l=0}^{j} f_l T^l \right) = \sum_{l=0}^{j} \phi(f_l) T^l.
\]

It follows that \( \phi \) is uniquely determined on \( A[[T]] \) by its restriction to \( A \), considered as a submodule of \( A[[T]] \) as a module over \( k \). We have also seen that every homomorphism from \( A \) into \( B[[T]] \), as modules over \( k \), can be extended to a homomorphism from \( A[[T]] \) into \( B[[T]] \) as modules over \( k[[T]] \), as in Sections 4.4 and 4.6.

This gives a natural isomorphism between \( \text{Hom}_k(A, B[[T]]) \) and \( \text{Hom}_{k[[T]]}(A[[T]], B[[T]]) \),

\[
\text{Hom}_{k[[T]]}(A[[T]], B[[T]]),
\]

\[
\text{Hom}_{k[[T]]}(A[[T]], B[[T]]),
\]

\[
\text{Hom}_{k[[T]]}(A[[T]], B[[T]])
\]
as modules over $k$. We may also consider $\text{Hom}_k(A, B[[T]])$ as a module over $k[[T]]$, because it is a submodule of the space of all $B[[T]]$-valued functions on $A$, as a module over $k[[T]]$. It is easy to see that the isomorphism between $\text{Hom}_k(A, B[[T]])$ and (4.8.5) is linear over $k[[T]]$.

Let $C$ be another module over $k$, and let $\beta$ be a mapping from $A[T] \times B[T]$ into $C[T]$. If $\beta$ is bilinear over $k$ and

$$\beta(f(T), g(T)) = \beta(f(T), g(T), T) = \beta(f(T), g(T))T$$

for every $f(T) \in A[T]$ and $g(T) \in B[T]$, then

$$\beta(f(T)^r, g(T)^m) = \beta(f(T), g(T))T^{r + m}$$

for all $f(T) \in A[T]$, $g(T) \in B[T]$, and nonnegative integers $r, m$, and hence $\beta$ is bilinear over $k[T]$. Conversely, if $\beta$ is bilinear over $k[T]$, then $\beta$ is bilinear over $k$ and satisfies (4.8.6). It is easy to see that $\beta$ is uniquely determined on $A[T] \times B[T]$ by its restriction to $A \times B$ in this situation. We have seen that every mapping from $A \times B$ into $C[T]$ that is bilinear over $k$ can be extended to a mapping from $A[T] \times B[T]$ into $C[T]$ that is bilinear over $k[T]$, as in Sections 4.5 and 4.6.

Let $\beta$ be a mapping from $A[[T]] \times B[[T]]$ into $C[[T]]$. Suppose that $\beta$ is bilinear over $k$ and satisfies (4.8.6) for every $f(T) \in A[[T]]$ and $g(T) \in B[[T]]$, which implies that (4.8.7) holds for all $f(T) \in A[[T]]$, $g(T) \in B[[T]]$, and nonnegative integers $r, m$. If $f(T) \in A[[T]]$, $g(T) \in B[[T]]$ vanish to orders $r_1, r_2$ for some nonnegative integers $r_1, r_2$, respectively, then it follows that $\beta(f(T), g(T))$ vanishes to order $r_1 + r_2$. One can use this to verify that $\beta$ is bilinear over $k[[T]]$.

Conversely, if $\beta$ is bilinear over $k[[T]]$, then $\beta$ is bilinear over $k[T]$, and hence $\beta$ is bilinear over $k$ and satisfies (4.8.6) for every $f(T) \in A[[T]]$ and $g(T) \in B[[T]]$.

Suppose that $\beta$ is bilinear over $k$ and satisfies (4.8.6) for every $f(T) \in A[[T]]$ and $g(T) \in B[[T]]$ again. Let $f(T) = \sum_{j=0}^{\infty} f_j T^j \in A[[T]]$, and $g(T) = \sum_{i=0}^{\infty} g_i T^i \in B[[T]]$ be given. If $r$ is any nonnegative integer, then the total coefficient of $T^r$ in $\beta(f(T), g(T))$ is the same as for

$$\beta \left( \sum_{j=0}^{r} f_j T^j, \sum_{i=0}^{r} g_i T^i \right).$$

This implies that $\beta$ is uniquely determined on $A[[T]] \times B[[T]]$ by its restriction to $A \times B$. We have seen that every mapping from $A \times B$ into $C[[T]]$ that is bilinear over $k$ can be extended to a mapping from $A[[T]] \times B[[T]]$ into $C[[T]]$ that is bilinear over $k[[T]]$, as in Sections 4.5 and 4.6.

### 4.9 Formal power series and homomorphisms

Let $k$ be a commutative ring with a multiplicative identity element again, let $A$, $B$ be modules over $k$, and let $T$ be an indeterminate. Remember that the space $\text{Hom}_k(A, B)$ of module homomorphisms from $A$ into $B$ is a module over $k$ with
4.9. FORMAL POWER SERIES AND HOMOMORPHISMS

respect to pointwise addition and scalar multiplication, as in Section 2.1. Thus the corresponding spaces \((\text{Hom}_k(A, B))[T]\) and \((\text{Hom}_k(A, B))[[T]]\) of formal polynomials and power series in \(T\) with coefficients in \(\text{Hom}_k(A, B)\) can be defined as in Section 4.3. More precisely, \((\text{Hom}_k(A, B))[T]\) and \((\text{Hom}_k(A, B))[[T]]\) may be considered as modules over \(k[T]\) and \(k[[T]]\), respectively, as in Section 4.6. Let us see how elements of these modules are related to homomorphisms from \(A\) into \(B[T]\) and \(B[[T]]\), respectively.

Let

\[
\phi(T) = \sum_{j=0}^{\infty} \phi_j T^j
\]

be a formal power series in \(T\) with coefficients in \(\text{Hom}_k(A, B)\), so that \(\phi_j\) is a homomorphism from \(A\) into \(B\) for every \(j \geq 0\). If \(a \in A\), then

\[
(\phi(T))(a) = \sum_{j=0}^{\infty} \phi_j(a) T^j
\]

defines a formal power series in \(T\) with coefficients in \(B\), and the mapping from \(a \in A\) to (4.9.2) is a homomorphism from \(A\) into \(B[T]\), as modules over \(k\). Conversely, every homomorphism from \(A\) into \(B[[T]]\) as modules over \(k\) corresponds to a sequence \(\{\phi_j\}_{j=0}^{\infty}\) of homomorphisms from \(A\) into \(B\) in this way, and hence to an element of \((\text{Hom}_k(A, B))[[T]]\). This defines a natural isomorphism between \(\text{Hom}_k(A, B[[T]])\) and \((\text{Hom}_k(A, B))[[T]]\), as modules over \(k\). This isomorphism is linear over \(k[[T]]\) as well, where \(\text{Hom}_k(A, B[[T]])\) is considered as a module over \(k[[T]]\), as in the previous section.

Similarly, if

\[
\phi(T) = \sum_{j=0}^{n} \phi_j T^j
\]

is a formal polynomial in \(T\) with coefficients in \(\text{Hom}_k(A, B)\), then

\[
(\phi(T))(a) = \sum_{j=0}^{n} \phi_j(a) T^j
\]

is a formal polynomial in \(T\) with coefficients in \(B\) for every \(a \in A\), and the mapping from \(a \in A\) to (4.9.4) is a homomorphism from \(A\) into \(B[T]\), as modules over \(k\). In the other direction, a homomorphism from \(A\) into \(B[T]\) as modules over \(k\) may be considered as a homomorphism from \(A\) into \(B[[T]]\), and corresponds to a sequence \(\{\phi_j\}_{j=0}^{\infty}\) of homomorphisms from \(A\) into \(B\) as in (4.9.2). The condition that this mapping sends \(A\) into \(B[T]\) means that for each \(a \in A\), we have that \(\phi_j(a) = 0\) for all but finitely many \(j\). In particular, this holds when \(\phi_j = 0\) for all but finitely many \(j\), as mappings from \(A\) into \(B\). If for every \(a \in A\) we have that \(\phi_j(a) = 0\) for all but finitely many \(j\), and if \(A\) is finitely generated as a module over \(k\), then \(\phi_j = 0\) for all but finitely many \(j\).

Consider the mapping from (4.9.3) to (4.9.4), as a module homomorphism from \(A\) into \(B[T]\). This defines a natural homomorphism from \((\text{Hom}_k(A, B))[T]\)
into $\text{Hom}_k(A, B[[T]])$, as modules over $k$. It is easy to see that this mapping is injective. This mapping is also linear over $k[T]$, where $\text{Hom}_k(A, B[[T]])$ is considered as a module over $k[T]$, as in the previous section. If $A$ is finitely generated as a module over $k$, then this mapping is surjective, as in the preceding paragraph.

### 4.10 Extensions and compositions

Let us continue with the same notation and hypotheses as in the preceding section. Let $a(T) = \sum_{m=0}^{\infty} a_m T^m$ be a formal power series in $T$ with coefficients in $A$. If $\phi(T)$ is a formal power series in $T$ with coefficients in $\text{Hom}_k(A, B)$ as in (4.9.1), then

$$(4.10.1) \quad (\phi(T))(a_m) = \sum_{j=0}^{\infty} \phi_j(a_m) T^j$$

defines a formal power series in $T$ with coefficients in $B$ for every nonnegative integer $m$, as in (4.9.2). Using this, we can define

$$(4.10.2) \quad (\phi(T))(a(T)) = \sum_{m=0}^{\infty} (\phi(T))(a_m) T^m$$

as a formal power series in $T$ with coefficients in $B$, as in Section 4.4. This corresponds to extending the homomorphism from $A$ into $B[[T]]$ associated to $\phi(T)$ to a homomorphism from $A[[T]]$ into $B[[T]]$, as modules over $k[[T]]$, as before. Of course, if $a(T)$ and $\phi(T)$ are formal polynomials in $T$ with coefficients in $A$ and $\text{Hom}_k(A, B)$, respectively, then (4.10.1) is a formal polynomial in $T$ with coefficients in $B$ for each $m \geq 0$, and (4.10.2) is a formal polynomial in $T$ with coefficients in $B$ too. This corresponds to extending the homomorphism from $A$ into $B[[T]]$ associated to $\phi(T)$ to a homomorphism from $A[T]$ into $B[T]$, as modules over $k[T]$.

Alternatively, put

$$(4.10.3) \quad E(\phi, a) = \phi(a)$$

for every $\phi \in \text{Hom}_k(A, B)$ and $a \in A$, which is the natural evaluation mapping

$$(4.10.4) \quad \text{from } \text{Hom}_k(A, B) \times A \text{ into } B.$$

This mapping is bilinear over $k$, and can be extended to a mapping

$$(4.10.5) \quad \text{from } (\text{Hom}_k(A, B))[[T]] \times A[[T]] \text{ into } B[[T]],$$

as in Section 4.5. More precisely, let $\phi(T) \in (\text{Hom}_k(A, B))[[T]]$ and $a(T)$ in $A[[T]]$ be given as before, and put

$$(4.10.6) \quad E_n(\phi(T), a(T)) = \sum_{j=0}^{n} \phi_j(a_{n-j})$$
for each nonnegative integer $n$. This is an element of $B$ for every $n \geq 0$, so that

$$(4.10.7) \quad E(\phi(T), a(T)) = \sum_{n=0}^{\infty} E_n(\phi(T), a(T)) T^n$$

is an element of $B[[T]]$. It is easy to see that this is the same as (4.10.2). If $\phi(T) \in (\text{Hom}_k(A, B))[[T]]$ and $a(T) \in A[T]$, then (4.10.6) is equal to 0 for all but finitely many $n \geq 0$, so that (4.10.6) is an element of $B[T]$. This corresponds to extending (4.10.3) to a mapping

$$(4.10.8) \quad \text{from } (\text{Hom}_k(A, B))[[T]] \times A[T] \text{ into } B[T],$$

as before.

Let $C$ be another module over $k$, let $\phi(T)$ be a formal power series in $T$ with coefficients in $\text{Hom}_k(A, B)$ as in (4.9.1), and let

$$(4.10.9) \quad \psi(T) = \sum_{l=0}^{\infty} \psi_l T^l$$

be a formal power series in $T$ with coefficients in $\text{Hom}_k(B, C)$. Note that the composition $\psi_l \circ \phi_j$ of $\phi_j$ and $\psi_l$ defines a module homomorphism from $A$ into $C$ for all $j, l \geq 0$. Thus

$$(4.10.10) \quad (\psi(T) \circ \phi(T))_n = \sum_{l=0}^{n} \psi_l \circ \phi_{n-l}$$

is an element of $\text{Hom}_k(A, C)$ for every nonnegative integer $n$. Put

$$(4.10.11) \quad \psi(T) \circ \phi(T) = \sum_{n=0}^{\infty} (\psi(T) \circ \phi(T))_n T^n,$$

which defines a formal power series in $T$ with coefficients in $\text{Hom}_k(A, C)$. If $\phi(T)$ and $\psi(T)$ are formal polynomials in $T$, then (4.10.10) is equal to 0 for all but finitely many $n \geq 0$, so that (4.10.11) is a formal polynomial in $T$ as well.

The composition of module homomorphisms from $A$ into $B$ and from $B$ into $C$ defines a mapping

$$(4.10.12) \quad \text{from } \text{Hom}_k(A, B) \times \text{Hom}_k(B, C) \text{ into } \text{Hom}_k(A, C)$$

that is bilinear over $k$. The definition of (4.10.11) corresponds to the extension of this bilinear mapping to mappings

$$(4.10.13) \quad \text{from } (\text{Hom}_k(A, B))[[T]] \times (\text{Hom}_k(B, C))[[T]] \text{ into } (\text{Hom}_k(A, C))[[T]]$$

and

$$(4.10.14) \quad \text{from } (\text{Hom}_k(A, B))[T] \times (\text{Hom}_k(B, C))[T] \text{ into } (\text{Hom}_k(A, C))[T].$$
as in Section 4.5.

As before, \( \phi(T) \in (\text{Hom}_k(A,B))[T] \) corresponds to a homomorphism from \( A \) into \( B[[T]] \) as modules over \( k \), which can be extended to a homomorphism from \( A[[T]] \) into \( B[[T]] \) as modules over \( k[[T]] \). Similarly, \( \psi(T) \) corresponds to a homomorphism from \( B[[T]] \) into \( C[[T]] \) as modules over \( k[[T]] \), and (4.10.11) corresponds to a homomorphism from \( A[[T]] \) into \( C[[T]] \) as modules over \( k[[T]] \).

One can check that the homomorphism corresponding to (4.10.11) is the same as the composition of the homomorphisms corresponding to \( \phi(T) \) and \( \psi(T) \).

If \( \phi(T) \in (\text{Hom}_k(A,B))[T] \) and \( \psi(T) \in (\text{Hom}_k(B,C))[T] \), then (4.10.11) is an element of \( (\text{Hom}_k(A,C))[T] \), as before. In this case, (4.10.11) corresponds to a homomorphism from \( A[T] \) into \( C[T] \), as modules over \( k[T] \), which is the composition of the homomorphisms from \( A[T] \) into \( B[T] \) and from \( B[T] \) into \( C[T] \) corresponding to \( \phi(T) \) and \( \psi(T) \), respectively.

### 4.11 Two-step extensions

Let \( k \) be a commutative ring with a multiplicative identity element, let \( A, B, C \) be modules over \( k \), and let \( T \) be an indeterminate. If \( \beta \) is a mapping from \( A \times B \) into \( C[[T]] \) that is bilinear over \( k \), then \( \beta \) can be extended to a mapping from \( A[[T]] \times B[[T]] \) into \( C[[T]] \) that is bilinear over \( k[[T]] \), as in Sections 4.5 and 4.6. One can also look at this in terms of extending module homomorphisms, as in Sections 4.4 and 4.6, in two steps. More precisely, if \( a \in A \), then \( \beta(a,b) \) defines a mapping from \( B \) into \( C[[T]] \), as a function of \( b \), that is linear over \( k \). This can be extended to a mapping from \( B[[T]] \) into \( C[[T]] \) that is linear over \( k[[T]] \), as before. If \( b(T) \in B[[T]] \), then we can use the extension just mentioned to get \( \beta(a,b(T)) \) as a function of \( a \in A \) with values in \( C[[T]] \) that is linear over \( k \). This can be extended to a mapping from \( A[[T]] \) into \( C[[T]] \) that is linear over \( k[[T]] \), which extends \( \beta \) to a mapping from \( A[[T]] \times B[[T]] \) into \( C[[T]] \) that is bilinear over \( k[[T]] \). Similarly, if \( \beta \) is a mapping from \( A \times B \) into \( C[T] \) that is bilinear over \( k \), then \( \beta \) can be extended to a mapping from \( A[T] \times B[T] \) into \( C[T] \) that is bilinear over \( k[T] \) in two steps.

Let \( \beta \) be a mapping from \( A \times B \) into \( C[[T]] \) that is bilinear over \( k \) again, and put

\[
(4.11.1) \quad \rho_a(b) = \beta(a,b)
\]

for every \( a \in A \) and \( b \in B \). If \( a \in A \), then \( \rho_a \) defines a mapping from \( B \) into \( C[[T]] \) that is linear over \( k \), which is to say that \( \rho_a \in \text{Hom}_k(B,C[[T]]) \). Thus

\[
(4.11.2) \quad a \mapsto \rho_a
\]

defines a mapping from \( A \) into \( \text{Hom}_k(B,C[[T]]) \), which is linear over \( k \). Remember that there are natural isomorphisms between \( \text{Hom}_k(B,C[[T]]) \) and each of \( \text{Hom}_k((B[[T]],C[[T]]) \) and \( (\text{Hom}_k(B,C))[[T]] \), as modules over \( k[[T]] \), as in Sections 4.8 and 4.9. The isomorphism with \( \text{Hom}_k((B[[T]], C[[T]]) \) gives the extension of \( \rho_a \) to \( B[[T]] \), and the isomorphism with \( (\text{Hom}_k(B,C))[[T]] \) can be used to extend (4.11.2) to \( A[[T]] \).
4.12 Extending algebra homomorphisms

Let $k$ be a commutative ring with a multiplicative identity element, let $T$ be an indeterminate, and let $A$, $B$ be algebras over $k$ in the strict sense. As in Section 4.6, multiplication on $A$ and $B$ can be extended to $A[[T]]$ and $B[[T]]$, respectively, so that they become algebras over $k[[T]]$ in the strict sense. Moreover, $A[[T]]$ and $B[[T]]$ may be considered as algebras over $k$.

Let $\phi$ be a homomorphism from $A$ into $B[[T]]$, as modules over $k$. Thus $\phi$ can be expressed as

$$\phi(a) = \sum_{j=0}^{\infty} \phi_j(a) T^j$$

for each $a \in A$, where $\phi_j$ is a homomorphism from $A$ into $B$, as modules over $k$, for every nonnegative integer $j$. In order for $\phi$ to be a homomorphism from $A$ into $B[[T]]$, as algebras over $k$, we need to have that

$$\phi(a a') = \phi(a) \phi(a')$$

for every $a, a' \in A$. Of course,

$$\phi(a a') = \sum_{n=0}^{\infty} \phi_n(a a') T^n,$$

and

$$\phi(a) \phi(a') = \left( \sum_{j=0}^{\infty} \phi_j(a) T^j \right) \left( \sum_{l=0}^{\infty} \phi_l(a') T^l \right) = \sum_{n=0}^{\infty} \left( \sum_{j=0}^{n} \phi_j(a) \phi_{n-j}(a') \right) T^n.$$
It follows that (4.12.2) holds if and only if

\[ \phi_n(a a') = \sum_{j=0}^{n} \phi_j(a) \phi_{n-j}(a') \] (4.12.5)

for every nonnegative integer \( n \).

As in Section 4.4, we can extend \( \phi \) to a module homomorphism from \( A[[T]] \) into \( B[[T]] \) by putting

\[ \phi(f(T)) = \sum_{j=0}^{\infty} \phi(f_j) T^j \] (4.12.6)

for every \( f(T) = \sum_{j=0}^{\infty} f_j T^j \in A[[T]] \). If \( g(T) = \sum_{l=0}^{\infty} g_l T^l \) is another element of \( A[[T]] \), then \( h(T) = f(T)g(T) \) is defined by putting \( h(T) = \sum_{n=0}^{\infty} h_n T^n \), where

\[ h_n = \sum_{j=0}^{n} f_j g_{n-j} \] (4.12.7)

for every \( n \geq 0 \), as in Section 4.6. Thus

\[ \phi(f(T)g(T)) = \phi(h(T)) = \sum_{n=0}^{\infty} \phi(h_n) T^n \] (4.12.8)

where the sum on the right is defined as an element of \( B[[T]] \) as in Section 4.4 again. If \( \phi \) is an algebra homomorphism from \( A \) into \( B[[T]] \), then

\[ \phi(h_n) = \sum_{j=0}^{n} \phi(f_j) g_{n-j} = \sum_{j=0}^{n} \phi(f_j) \phi(g_{n-j}) \] (4.12.9)

for each \( n \geq 0 \). This implies that

\[ \phi(f(T)g(T)) = \sum_{n=0}^{\infty} \left( \sum_{j=0}^{n} \phi(f_j) \phi(g_{n-j}) \right) T^n = \phi(f(T)) \phi(g(T)). \] (4.12.10)

More precisely, this follows from the definition of multiplication on \( B[[T]] \) when \( \phi \) maps \( A \) into \( B \). Otherwise, if \( \phi(f_j) \) and \( \phi(g_l) \) are elements of \( B[[T]] \), then the second step in (4.12.10) can be obtained as in (4.5.6). This shows that the extension of \( \phi \) to a module homomorphism from \( A[[T]] \) into \( B[[T]] \) defined in (4.12.6) is an algebra homomorphism in this case.

Similarly, multiplication on \( A \) and \( B \) can be extended to \( A[T] \) and \( B[T] \), respectively, so that they become algebras in the strict sense over \( k[T] \), as in Section 4.6. If \( \phi \) is a homomorphism from \( A \) into \( B[T] \), as modules over \( k \), then we can extend \( \phi \) to a module homomorphism from \( A[T] \) into \( B[T] \) as in (4.12.6). If \( \phi \) is an algebra homomorphism from \( A \) into \( B[T] \), then this extension is an algebra homomorphism from \( A[T] \) into \( B[T] \), as before. Of course, there are analogous statements for opposite algebra homomorphisms.
4.13 Some remarks about commutators

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( A \) be an associative algebra over \( k \) with a multiplicative identity element \( e \). Also let \( T \) be an indeterminate, and let us identify \( e \) with the corresponding formal power series in \( T \) with coefficients in \( A \) in the usual way, which is the multiplicative identity element in \( A[[T]] \). Suppose that \( a(T) = \sum_{j=0}^{\infty} a_j T^j \) and \( b(T) = \sum_{j=0}^{\infty} b_j T^j \) are elements of \( A[[T]] \) with \( a_0 = b_0 = e \). Note that \( a(T) \) and \( b(T) \) are invertible in \( A[[T]] \), as in Section 4.7.

Let \( \alpha(T) = a(T) - e = \sum_{j=1}^{\infty} a_j T^j \) and \( \beta(T) = b(T) - e = \sum_{j=1}^{\infty} b_j T^j \), so that \( a(T) = e + \alpha(T) \) and \( b(T) = e + \beta(T) \). Thus

\[
(4.13.1) \quad a(T) b(T) = e + \alpha(T) + \beta(T) + \alpha(T) \beta(T).
\]

As in Section 4.7,

\[
(4.13.2) \quad a(T)^{-1} = \sum_{l=0}^{\infty} (-\alpha(T))^l = e + \sum_{l=1}^{\infty} (-\alpha(T))^l
\]

and

\[
(4.13.3) \quad b(T)^{-1} = \sum_{l=0}^{\infty} (-\beta(T))^l = e + \sum_{l=1}^{\infty} (-\beta(T))^l.
\]

If \( n \in \mathbb{Z}_+ \), then we let \( O(T^n) \) refer to any element of \( A[[T]] \) that vanishes to order \( n - 1 \), which means that it can be expressed as an element of \( A[[T]] \) times \( T^n \). Observe that

\[
(4.13.4) \quad a(T) b(T) a(T)^{-1} b(T)^{-1} = e + O(T),
\]

because \( a(T), b(T), a(T)^{-1}, b(T)^{-1} = e + O(T) \). More precisely,

\[
(4.13.5) \quad a(T) b(T) a(T)^{-1} b(T)^{-1} = e + O(T^2),
\]

because

\[
(4.13.6) \quad a(T) = e + a_1 T + O(T^2), \quad b(T) = e + b_1 T + O(T^2),
\]

and

\[
(4.13.7) \quad a(T)^{-1} = e - a_1 T + O(T^2), \quad b(T)^{-1} = e - b_1 T + O(T^2).
\]

We also have that

\[
(4.13.8) \quad a(T) = e + a_1 T + a_2 T^2 + O(T^3), \quad b(T) = e + b_1 T + b_2 T^2 + O(T^3),
\]

and

\[
(4.13.9) \quad a(T) b(T) = e + a_1 T + b_1 T + a_2 T^2 + b_2 T^2 + a_1 b_1 T^2 + O(T^3).
\]

Using (4.13.2), we get that

\[
(4.13.10) \quad a(T)^{-1} = e - \alpha(T)^2 + O(T^3)
\]

\[
= e - a_1 T - a_2 T^2 + a_1^2 T^2 + O(T^3).
\]
Similarly,
\[(4.13.11) \quad (T)^{-1} = e - b_1 T - b_2 T^2 + b_3 T^3 + O(T^5).\]
It follows that
\[(4.13.12) \quad a(T)^{-1} b(T)^{-1} = e - a_1 T - b_1 T - a_2 T^2 - b_2 T^2 + a_3 T^3 + O(T^4).\]

Combining (4.13.9) and (4.13.12), it is not difficult to verify that
\[(4.13.13) \quad a(T) b(T) a(T)^{-1} = e + a_1 b_1 T^2 - b_1 a_1 T^2 + O(T^5).\]

4.14 Formal power series and involutions

Let \(k\) be a commutative ring with a multiplicative identity element, and let \(T\) be an indeterminate. Also let \(A\) be an algebra over \(k\) in the strict sense, and let \(x \mapsto x^*\) be an opposite algebra automorphism on \(A\). If \(k = \mathbb{C}\), then we may wish to consider opposite algebra automorphisms that are conjugate-linear, as usual. If \(a(T) = \sum_{j=0}^{\infty} a_j T^j \in A[[T]],\) then
\[(4.14.1) \quad a(T)^* = \sum_{j=0}^{\infty} a_j^* T^j\]
defines an element of \(A[[T]],\) which is in \(A[T]\) when \(a(T) \in A[T].\) Of course,
\[(4.14.2) \quad a(T) \mapsto a(T)^*\]
defines an opposite algebra automorphism on \(A[[T]].\) Note that \(a(T)\) is self-adjoint with respect to (4.14.2) if and only if \(a_j\) is self-adjoint for every \(j \geq 0,\) and similarly \(a(T)\) is anti-self-adjoint with respect to (4.14.2) if and only if \(a_j\) is anti-self-adjoint for every \(j \geq 0.\) If \(x \mapsto x^*\) is an involution on \(A,\) then (4.14.2) is an involution on \(A[[T]].\)

Suppose now that \(A\) is an associative algebra over \(k\) with a multiplicative identity element \(e.\) Remember that \(e^* = e,\) as in Section 2.6. Let \(y\) be an element of \(A,\) and put \(y_0 = y - e,\) so that \(y = e + y_0,\) \(y^* = e + y_0^*,\) and
\[(4.14.3) \quad y^* y = (e + y_0^*) (e + y_0) = e + y_0 + y_0^* + y_0 y_0.\]
Suppose that \(y_1, y_2 \in A\) satisfy
\[(4.14.4) \quad y_0 = y_1 + y_2, \quad y_1^* = y_1, \quad y_2^* = -y_2, \quad \text{and} \quad y_1 y_2 = y_2 y_1.\]
Under these conditions, \(y_0^* = y_1 - y_2,\) and
\[(4.14.5) \quad y^* y = e + 2 \cdot y_1 + (y_1 - y_2) (y_1 + y_2) = e + 2 \cdot y_1 + y_1^2 - y_2^2.\]
Let \( a(T) = \sum_{j=0}^{\infty} a_j T^j \) be an element of \( A[[T]] \) with \( a_0 = e \), and put \( \alpha(T) = a(T) - e \). As in (4.14.3), we have that
\[
(4.14.6) \quad a(T)^* a(T) = e + \alpha(T) + \alpha(T)^* a(T).
\]
In particular, a necessary condition for
\[
(4.14.7) \quad a(T)^* a(T) = e
\]
to hold is that
\[
(4.14.8) \quad a_1^* = -a_1.
\]
More precisely, (4.14.8) is equivalent to
\[
(4.14.9) \quad a(T)^* = a(T)^{-1},
\]
which also implies that \( a(T) \) commutes with \( a(T)^* \).

Suppose that \( \beta(T) = \sum_{j=0}^{\infty} \beta_j T^j, \gamma(T) = \sum_{j=0}^{\infty} \gamma_j T^j \in A[[T]] \) satisfy \( \beta_0 = \gamma_0 = 0 \), \( \alpha(T) = \beta(T) + \gamma(T) \),
\[
(4.14.10) \quad \beta(T)^* = \beta(T), \quad \gamma(T)^* = -\gamma(T), \quad \text{and} \quad \beta(T) \gamma(T) = \gamma(T) \beta(T).
\]
Under these conditions, we have that
\[
(4.14.11) \quad a(T)^* a(T) = e + 2 \cdot \beta(T) + \beta(T)^2 - \gamma(T)^2,
\]
as in (4.14.5). In this situation, (4.14.7) holds exactly when
\[
(4.14.12) \quad 2 \cdot \beta(T) = -\beta(T)^2 + \gamma(T)^2.
\]
Let us suppose from now on in this section that \( 1+1 \) has a multiplicative inverse in \( k \). It is easy to see that \( \beta(T) \) is uniquely determined by \( \gamma(T) \) and (4.14.12).

More precisely, if \( \gamma(T) \in A[[T]] \) satisfies \( \gamma_0 = 0 \), then there is a unique \( \beta(T) \in A[[T]] \) that satisfies \( \beta_0 = 0 \) and (4.14.12). Indeed, for each \( j \in \mathbb{Z}_+ \), one can use (4.14.12) to get \( \beta_l \) from \( \beta_l \) with \( l < j \) and the coefficients of \( \gamma(T) \). One can in fact approximate \( \beta \) in terms of polynomials in \( \gamma(T)^2 \) with coefficients in \( k \), using (4.14.12). This implies that \( \beta(T) \) commutes with \( \gamma(T) \), which could also be verified more directly. If \( \gamma(T)^* = -\gamma(T) \), as in (4.14.10), then
\[
(4.14.13) \quad (\gamma(T)^2)^* = (\gamma(T)^*)^2 = (-\gamma(T))^2 = \gamma(T)^2.
\]
In this case, one can check that \( \beta(T)^* = \beta(T) \). If we take \( \alpha(T) = \beta(T) + \gamma(T) \) and \( a(T) = e + \alpha(T) \), as before, then (4.14.11) holds, which implies (4.14.7), by (4.14.12).
Chapter 5

Some related notions

5.1 Adjoining nilpotent elements

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be a module over $k$. If $n \in \mathbb{Z}_+$, then we would like to define $A[\epsilon_n]$ as a module over $k$, where $\epsilon_n$ is an additional element that is considered to satisfy

$$\epsilon_n^{n+1} = 0. \tag{5.1.1}$$

The elements of $A[\epsilon_n]$ may be expressed as formal sums of the form

$$a = a_0 + a_1 \epsilon_n + \cdots + a_n \epsilon_n^n, \tag{5.1.2}$$

where $a_0, a_1, \ldots, a_n \in A$. Addition and scalar multiplication on $A[\epsilon_n]$ are defined termwise, so that $A[\epsilon_n]$ becomes a module over $k$ that contains $A$ as a submodule. More precisely, $A[\epsilon_n]$ is isomorphic to the direct sum of $n+1$ copies of $A$, as a module over $k$. If $a \in A[\epsilon_n]$ is as in (5.1.2), then

$$a \epsilon_n = a_0 \epsilon_n + \cdots + a_{n-1} \epsilon_n^n \tag{5.1.3}$$

defines an element of $A[\epsilon_n]$ as well. This defines

$$a \mapsto a \epsilon_n \tag{5.1.4}$$

as a module homomorphism from $A[\epsilon_n]$ into itself. It is sometimes convenient to consider the $a_0$ term on the right side of (5.1.2) as being $a_0 \epsilon_n^0$, so that multiplication by $\epsilon_n^n$ corresponds to the identity mapping on $A[\epsilon_n]$.

Let $B$ be another module over $k$, so that $B[\epsilon_n]$ can be defined as in the previous paragraph. Also let $\phi$ be a homomorphism from $A$ into $B[\epsilon_n]$, as modules over $k$. If $a \in A[\epsilon_n]$ is as in (5.1.2), then

$$\phi(a) = \phi(a_0) + \phi(a_1) \epsilon_n + \cdots + \phi(a_n) \epsilon_n^n \tag{5.1.5}$$

defines an element of $B[\epsilon_n]$. This defines an extension of $\phi$ to a homomorphism from $A[\epsilon_n]$ into $B[\epsilon_n]$, as modules over $k$. It is easy to see that

$$\phi(a \epsilon_n) = \phi(a) \epsilon_n \tag{5.1.6}$$

100
5.2. ADJOINING NILPOTENT ELEMENTS TO ALGEBRAS

for every \( a \in A[\epsilon_n] \), and that this extension of \( \phi \) to \( A[\epsilon_n] \) is uniquely determined by these properties.

Let \( C \) be another module over \( k \), so that \( C[\epsilon_n] \) can be defined as before. Also let \( \beta \) be a mapping from \( A \times B \) into \( C[\epsilon_n] \) that is bilinear over \( k \). We can extend \( \beta \) to a mapping from \( A[\epsilon_n] \times B[\epsilon_n] \) into \( C[\epsilon_n] \), as follows. Let \( a \in A[\epsilon_n] \) be as in (5.1.2), and let

\[
b = b_0 + b_1 \epsilon_n + \cdots + b_n \epsilon_n^n
\]

be an element of \( B[\epsilon_n] \), with \( b_0, b_1, \ldots, b_n \in B \). Thus

\[
\beta(a, b) = \sum_{j=0}^{n} \sum_{l=0}^{n} \beta(a_j, b_l) \epsilon_n^{j+l}
\]

defines an element of \( C[\epsilon_n] \), which extends \( \beta \) to a mapping from \( A[\epsilon_n] \times B[\epsilon_n] \) into \( C[\epsilon_n] \) that is bilinear over \( k \). One can check that

\[
\beta(a \epsilon_n, b) = \beta(a, b \epsilon_n) = \beta(a, b) \epsilon_n
\]

for every \( a \in A[\epsilon_n] \) and \( b \in B[\epsilon_n] \), and that this extension of \( \beta \) to \( A[\epsilon_n] \times B[\epsilon_n] \) is uniquely determined by the properties. One can also look at this in terms of extending \( \beta \) in each variable separately, as in the previous paragraph.

Suppose now that \( A = B \), so that we start with a mapping \( \beta \) from \( A \times A \) into \( C[\epsilon_n] \). If \( \beta \) is symmetric or antisymmetric on \( A \times A \), then the extension of \( \beta \) to \( A[\epsilon_n] \times A[\epsilon_n] \) into \( C[\epsilon_n] \) as in the preceding paragraph has the same property. Similarly, if

\[
\beta(a, a) = 0
\]

for every \( a \in A \), then (5.1.10) holds for every \( a \in A[\epsilon_n] \). To see this, let \( a \in A[\epsilon_n] \) be given as in (5.1.2), so that

\[
\beta(a, a) = \sum_{j=0}^{n} \sum_{l=0}^{n} \beta(a_j, a_l) \epsilon_n^{j+l}
\]

as in (5.1.8). By hypothesis, \( \beta(a_j, a_j) = 0 \) for each \( j \). Remember that \( \beta \) is antisymmetric on \( A \times A \), as in Section 2.1. This implies that

\[
\beta(a_j, a_l) + \beta(a_l, a_j) = 0
\]

when \( j \neq l \), which can be used to get that (5.1.11) is equal to 0.

5.2 Adjoining nilpotent elements to algebras

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( A \) be an algebra over \( k \) in the strict sense, where multiplication of \( a, b \in A \) is expressed as \( a \cdot b \). Also let \( n \) be a positive integer, and let \( \epsilon_n \) be as in the previous
section, so that \( A[\epsilon_n] \) can be defined as a module over \( k \) as before. If \( a, b \in A[\epsilon_n] \) are as in (5.1.2) and (5.1.7), respectively, then

\[
(5.2.1) \quad ab = \sum_{j=0}^{n} \sum_{l=0}^{n} a_{j} b_{l} \epsilon_{n}^{j+l}
\]
defines an element of \( A[\epsilon_n] \) too. This extends multiplication on \( A \) to a mapping from \( A[\epsilon_n] \times A[\epsilon_n] \) into \( A[\epsilon_n] \) that is bilinear over \( k \), as in the previous section. Thus \( A[\epsilon_n] \) becomes an algebra over \( k \) in the strict sense as well.

If multiplication on \( A \) is commutative, then this extension of multiplication to \( A[\epsilon_n] \) is commutative too. If multiplication on \( A \) is associative, then one can verify that multiplication on \( A[\epsilon_n] \) is associative as well. If \( A \) has a multiplicative identity element \( e \), then \( e \) is also the multiplicative identity element in \( A[\epsilon_n] \).

We can apply this to \( A = k \), to get \( k[\epsilon_n] \) as a commutative associative algebra over \( k \) with a multiplicative identity element.

Let \( A \) be a module over \( k \), so that \( A[\epsilon_k] \) can be defined as a module over \( k \) as in the previous section. In fact, \( A[\epsilon_n] \) may be considered as a module over \( k[\epsilon_n] \). More precisely, let \( a \in k[\epsilon_n] \) be as in (5.1.2), with \( a_0, a_1, \ldots, a_n \in k \), and let \( b \in A[\epsilon_n] \) be as in (5.1.7), with \( b_0, b_1, \ldots, b_n \in A \). Under these conditions, \( ab \) can be defined as an element of \( A[\epsilon_n] \) as in (5.2.1), where \( a_{j} b_{l} \) is defined as an element of \( A \) using scalar multiplication with respect to \( k \). This is the same as extending scalar multiplication on \( A \) as a bilinear mapping from \( k \times A \) into \( A \) to a bilinear mapping from \( k[\epsilon_n] \times A[\epsilon_n] \) into \( A[\epsilon_n] \), as in the previous section.

Let \( B \) be another module over \( k \), so that \( B[\epsilon_n] \) can be defined as before. It is easy to see that a mapping \( \phi \) from \( A[\epsilon_n] \) into \( B[\epsilon_n] \) is linear over \( k[\epsilon_n] \) if and only if \( \phi \) is linear over \( k \) and (5.1.6) holds for every \( a \in A[\epsilon_n] \). Similarly, let \( C \) be a module over \( k \), and let \( C[\epsilon_n] \) be as before. One can check that a mapping \( \beta \) from \( A[\epsilon_n] \times B[\epsilon_n] \) into \( C[\epsilon_n] \) is bilinear over \( k[\epsilon_n] \) if and only if \( \beta \) is bilinear over \( k \) and satisfies (5.1.9) for every \( a \in A[\epsilon_n] \) and \( b \in B[\epsilon_n] \). If \( A \) is an algebra over \( k \) in the strict sense, then \( A[\epsilon_n] \) may be considered as an algebra over \( k[\epsilon_n] \) in the strict sense.

If \( (A, [\cdot, \cdot]_A) \) is a Lie algebra over \( k \), then \( [\cdot, \cdot]_A \) can be extended to \( A[\epsilon_n] \) as before. More precisely, if \( a, b \in A[\epsilon_n] \) are as in (5.1.2) and (5.1.7), respectively, then

\[
(5.2.2) \quad [a, b]_A = \sum_{j=0}^{n} \sum_{l=0}^{n} [a_{j}, b_{l}]_A \epsilon_{n}^{j+l}
\]
defines an element of \( A[\epsilon_n] \) as well. As in the previous section, \( [a, a]_A = 0 \) for every \( a \in A[\epsilon_n] \), because of the analogous property of \( [\cdot, \cdot]_A \) on \( A \). One can verify that (5.2.2) satisfies the Jacobi identity on \( A[\epsilon_n] \), using the Jacobi identity for \( [\cdot, \cdot]_A \) on \( A \). Thus \( A[\epsilon_n] \) is a Lie algebra over \( k \) with respect to (5.2.2), which may be considered as a Lie algebra over \( k[\epsilon_n] \).

Let \( A \) be a module over \( k \) again, and let \( T \) be an indeterminate, so that the space \( A[[T]] \) of formal power series in \( T \) with coefficients in \( A \) can be defined as in Section 4.3. If \( f(T) = \sum_{j=0}^{\infty} f_j \epsilon_t^j \) is an element of \( A[[T]] \), then

\[
(5.2.3) \quad f_0 + f_1 \epsilon_n + \cdots + f_n \epsilon_n^n
\]
5.3. **INVERTIBILITY IN $A[\epsilon_n]$**

defines an element of $A[\epsilon]$. This defines a homomorphism from $A[[T]]$ onto $A[\epsilon_n]$, as modules over $k$. This homomorphism also maps the space $A[T]$ of formal polynomials in $T$ with coefficients in $A$ onto $A[\epsilon_n]$. Note that multiplication by $T$ on $A[[T]]$ corresponds to multiplication by $\epsilon_n$ on $A[\epsilon_n]$ with respect to this homomorphism.

If $A$ is an algebra over $k$ in the strict sense, then one can check that the mapping from $f(T)$ in $A[[T]]$ to (5.2.3) defines a homomorphism from $A[[T]]$ onto $A[\epsilon_n]$ as algebras over $k$. In particular, we can apply this to $A = k$.

### 5.3 Invertibility in $A[\epsilon_n]$  

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an associative algebra over $k$ with a multiplicative identity element $e$. Also let $n$ be a positive integer, and let $A[\epsilon_n]$ be as in the previous two sections.

Remember that $a \in A[\epsilon_n]$ can be expressed as

\[ a = a_0 + a_1 \epsilon_n + \cdots + a_n \epsilon_n^m, \]

where $a_0, a_1, \ldots, a_n \in A$. The mapping

\[ a \mapsto a_0 \]

defines an algebra homomorphism from $A[\epsilon_n]$ onto $A$. If $e$ is invertible in $A[\epsilon_n]$, then it follows that $a_0$ is invertible in $A$.

Note that

\[ (e - a) \sum_{l=0}^{n} a^l = \left( \sum_{l=0}^{n} a^l \right) (e - a) = e - a^{n+1} \]

for every $a \in A[\epsilon_n]$, where $a^l$ is interpreted as being equal to $e$ when $l = 0$, as usual. If $a_0 = 0$, then $a^{n+1} = 0$, so that (5.3.3) reduces to

\[ (e - a) \sum_{l=0}^{n} a^l = \left( \sum_{l=0}^{n} a^l \right) (e - a) = e. \]

This means that $e - a$ is invertible in $A[\epsilon_n]$, with

\[ (e - a)^{-1} = \sum_{l=0}^{n} a^l. \]

If

\[ b = b_0 + b_1 \epsilon_n + \cdots + b_n \epsilon_n \]

is an element of $A[\epsilon_n]$, where $b_0$ is an invertible element of $A$, then $b$ can be expressed as $b_0 (e - a)$, where $a \in A[\epsilon_n]$ is as in (5.3.1), with $a_0 = 0$. It follows that $b$ is invertible in $A[\epsilon_n]$, because $e - a$ is invertible.
Suppose now that \( a, b \in A[\epsilon_n] \) are as in (5.3.1) and (5.3.6), respectively, with \( a_0 = b_0 = e \). Put \( \alpha = a - e \) and \( \beta = b - e \), so that \( a = e + \alpha, b = e + \beta \),

\[
ab = e + \alpha + \beta + \alpha \beta,
\]

and \( \alpha, \beta \) are multiples of \( \epsilon_n \). As in the preceding paragraph, \( a \) and \( b \) are invertible elements of \( A[\epsilon_n] \), with

\[
a^{-1} = \sum_{i=0}^{n} (-\alpha)^i = e + \sum_{i=1}^{n} (-\alpha)^i,
\]

and

\[
b^{-1} = \sum_{i=0}^{n} (-\beta)^i = e + \sum_{i=1}^{n} (-\beta)^i.
\]

We also have that

\[
aba^{-1}b^{-1} = e + a_1 b_1 \epsilon_n^2 - b_1 a_1 \epsilon_n^2 + O(\epsilon_n^3)
\]

\[
= e + \alpha \beta - \beta \alpha + O(\epsilon_n^3),
\]

where \( O(\epsilon_n^3) \) refers to any element of \( A[\epsilon_n] \) that is a multiple of \( \epsilon_n^3 \). This can be verified in the same way as in Section 4.13, or reduced to that situation, using the homomorphism from \( A[[T]] \) onto \( A[[\epsilon_n]] \) mentioned in the previous section.

If \( n = 1 \), then \( a = a_1 \epsilon_1, \beta = b_1 \epsilon_1 \), and (5.3.7) reduces to

\[
ab = e + \alpha + \beta = e + a_1 \epsilon_1 + b_1 \epsilon_1.
\]

In this case, (5.3.8) and (5.3.9) reduce to

\[
a^{-1} = e - \alpha = e - a_1 \epsilon_1
\]

and

\[
b^{-1} = e - \beta = e - b_1 \epsilon_1.
\]

Note that \( a \) and \( b \) commute in this situation, so that \( aba^{-1}b^{-1} = e \).

If \( n = 2 \), then \( a = a_1 \epsilon_2 + a_2 \epsilon_n^2, \beta = b_1 \epsilon_2 + b_2 \epsilon_n^2 \), and (5.3.7) reduces to

\[
ab = e + a_1 \epsilon_2 + a_2 \epsilon_n^2 + b_1 \epsilon_2 + b_2 \epsilon_n^2 + a_1 b_1 \epsilon_n^2.
\]

Similarly, (5.3.10) reduces to

\[
aba^{-1}b^{-1} = e + a_1 b_1 \epsilon_n^2 - b_1 a_1 \epsilon_n^2.
\]

### 5.4 Adjoining two nilpotent elements

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( A \) be a module over \( k \). We would like to define \( A[\epsilon_1, \eta_1] \) as a module over \( k \), where \( \epsilon_1 \) and \( \eta_1 \) are additional elements that are considered to satisfy

\[
\epsilon_1^2 = \eta_1^2 = 0, \quad \epsilon_1 \eta_1 = \eta_1 \epsilon_1.
\]
The elements of $A[\epsilon_1, \eta_1]$ can be expressed as formal sums of the form
\begin{equation}
(5.4.2) \quad a = a_{0,0} + a_{1,0} \epsilon_1 + a_{0,1} \eta_1 + a_{1,1} \epsilon_1 \eta_1,
\end{equation}
where $a_{0,0}, a_{1,0}, a_{0,1}, a_{1,1} \in A$. Addition and scalar multiplication on $A[\epsilon_1, \eta_1]$ are defined termwise, so that $A[\epsilon_1, \eta_1]$ becomes a module over $k$, which is isomorphic to the direct sum of four copies of $A$. By construction, $A[\epsilon_1, \eta_1]$ contains a copy of $A$, and in fact $A[\epsilon_1, \eta_1]$ contains copies of $A[\epsilon_1]$ and $A[\eta_1]$, which are defined as in Section 5.1. If $a \in A[\epsilon_1, \eta_1]$ is as in (5.4.2), then
\begin{equation}
(5.4.3) \quad a \epsilon_1 = a_{0,0} \epsilon_1 + a_{1,0} \epsilon_1 \eta_1, \quad a \eta_1 = a_{0,0} \eta_1 + a_{1,0} \epsilon_1 \eta_1
\end{equation}
define elements of $A[\epsilon_1, \eta_1]$. This defines
\begin{equation}
(5.4.4) \quad a \mapsto a \epsilon_1, \quad a \mapsto a \eta_1
\end{equation}
as module homomorphisms from $A[\epsilon_1, \eta_1]$ into itself. One can also look at $A[\epsilon_1, \eta_1]$ in terms of adjoining $\epsilon_1$ and $\eta_1$ separately, as before.

Let $B$ be another module over $k$, and let $\phi$ be a homomorphism from $A$ into $B[\epsilon_1, \eta_1]$, as modules over $k$, where $B[\epsilon_1, \eta_1]$ is defined as in the previous paragraph. If $a \in A[\epsilon_1, \eta_1]$ is as in (5.4.2), then
\begin{equation}
(5.4.5) \quad \phi(a) = \phi(a_{0,0}) + \phi(a_{1,0}) \epsilon_1 + \phi(a_{0,1}) \eta_1 + \phi(a_{1,1}) \epsilon_1 \eta_1
\end{equation}
defines an element of $B[\epsilon_1, \eta_1]$. This extends $\phi$ to a homomorphism from $A[\epsilon_1, \eta_1]$ into $B[\epsilon_1, \eta_1]$, as modules over $k$. This extension satisfies
\begin{equation}
(5.4.6) \quad \phi(a \epsilon_1) = \phi(a) \epsilon_1, \quad \phi(a \eta_1) = \phi(a) \eta_1
\end{equation}
for every $a \in A[\epsilon_1, \eta_1]$, and is uniquely determined by these properties.

Similarly, let $C$ be another module over $k$, and let $\beta$ be a mapping from $A \times B$ into $C[\epsilon_1, \eta_1]$ that is bilinear over $k$, where $C[\epsilon_1, \eta_1]$ is defined as before. One can extend $\beta$ to a mapping from $A[\epsilon_1, \eta_1] \times B[\epsilon_1, \eta_1]$ into $C[\epsilon_1, \eta_1]$ that is bilinear over $k$ and satisfies
\begin{equation}
(5.4.7) \quad \beta(a \epsilon_1, b) = \beta(a, b \epsilon_1) = \beta(a, b) \epsilon_1
\end{equation}
and
\begin{equation}
(5.4.8) \quad \beta(a \eta_1, b) = \beta(a, b \eta_1) = \beta(a, b) \eta_1
\end{equation}
for every $a \in A[\epsilon_1, \eta_1]$ and $b \in B[\epsilon_1, \eta_1]$.

Now let $A$ be an algebra over $k$ in the strict sense. Multiplication on $A$ can be extended to a mapping from $A[\epsilon_1, \eta_1] \times A[\epsilon_1, \eta_1]$ into $A[\epsilon_1, \eta_1]$ that is bilinear over $k$, as in the preceding paragraph, so that $A[\epsilon_1, \eta_1]$ becomes an algebra over $k$ in the strict sense too. If multiplication on $A$ is commutative, then multiplication on $A[\epsilon_1, \eta_1]$ is commutative as well. If multiplication on $A$ is associative, then one can check that multiplication on $A[\epsilon_1, \eta_1]$ is associative. If $A$ has a multiplicative identity element $e$, then $e$ is the multiplicative identity element in $A[\epsilon_1, \eta_1]$ too.

One can look at $A[\epsilon_1, \eta_1]$ as an algebra in terms adjoining $\epsilon_1$ and $\eta_1$ separately, as before. Thus these and other properties of $A[\epsilon_1, \eta_1]$ can be obtained from the remarks in Sections 5.1 and 5.2.
5.5 Invertibility in $A[\epsilon_1, \eta_1]$

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an associative algebra over $k$ with a multiplicative identity element $e$. Thus $A[\epsilon_1, \eta_1]$ is an associative algebra that contains $A$ as a subalgebra, as in the previous section, and $e$ is the multiplicative identity element of $A[\epsilon_1, \eta_1]$. As before, $a \in A[\epsilon_1, \eta_1]$ as in (5.4.2), with $a_{0,0}, a_{1,0}, a_{0,1}, a_{1,1} \in A$. The mapping

$$a \mapsto a_{0,0}$$

defines an algebra homomorphism from $A[\epsilon_1, \eta_1]$ onto $A$. If $a$ is invertible in $A[\epsilon_1, \eta_1]$, then $a_{0,0}$ is invertible in $A$.

As usual,

$$a \mapsto a_{0,0}, \quad b \mapsto b_{0,0}$$

for every $a \in A[\epsilon_1, \eta_1]$, where $a^l$ is interpreted as being $e$ when $l = 0$. If $a$ is as in (5.4.2), with $a_{0,0} = 0$, then $a^3 = 0$, and hence

$$a \mapsto a_{0,0}, \quad b \mapsto b_{0,0}$$

It follows that $a - e$ is invertible in $A[\epsilon_1, \eta_1]$, with

$$a - e = \sum_{l=0}^{2} a^l = 0.$$

If

$$b = b_{0,0} + b_{1,0} \epsilon_1 + b_{0,1} \eta_1 + b_{1,1} \epsilon_1 \eta_1$$

is an element of $A[\epsilon_1, \eta_1]$, where $b_{0,0}$ is an invertible element of $A$, then $b$ can be expressed as $b_{0,0} (e - a)$, where $a \in A[\epsilon_1, \eta_1]$ is as in (5.4.2), with $a_{0,0} = 0$. This implies that $b$ is invertible in $A[\epsilon_1, \eta_1]$, because $e - a$ is invertible.

Suppose that $a, b \in A[\epsilon_1, \eta_1]$ are as in (5.4.2) and (5.5.5), respectively, with $a_{0,0} = b_{0,0} = e$. Put

$$\alpha = a - e = a_{1,0} \epsilon_1 + a_{0,1} \eta_1 + a_{1,1} \epsilon_1 \eta_1$$

and

$$\beta = b - e = b_{1,0} \epsilon_1 + b_{0,1} \eta_1 + b_{1,1} \epsilon_1 \eta_1.$$

Thus $a = e + \alpha$, $b = e + \beta$, and

$$ab = e + \alpha + \beta + \alpha \beta = e + \alpha + \beta + a_{1,0} b_{0,1} \epsilon_1 \eta_1 + a_{0,1} b_{1,0} \epsilon_1 \eta_1.$$

We also have that $a$ and $b$ are invertible elements of $A[\epsilon_1, \eta_1]$, as in the preceding paragraph, with

$$a^{-1} = \sum_{l=0}^{2} (-\alpha)^l = e - \alpha + \alpha^2 = e - \alpha + 2 \cdot a_{1,0} a_{0,1} \epsilon_1 \eta_1$$
5.6. DIFFERENTIATION

and

\[(5.5.10)\quad b^{-1} = \sum_{l=0}^{2} (-\beta)^l = e - \beta + \beta^2 = e - \beta + 2 \cdot b_{1,0} b_{0,1} \epsilon_1 \eta_1.\]

Here \(2 \cdot x\) denotes \(x + x\) for any element \(x\) of \(A\) or \(A[\epsilon_1, \eta_1]\).

Suppose now that \(a_{0,1} = a_{1,1} = b_{1,0} = b_{1,1} = 0\), so that

\[(5.5.11)\quad a + a_{1,0} \epsilon_1, \quad b = e + b_{0,1} \eta_1\]

and \(a = a_{1,0} \epsilon_1, \quad \beta = b_{0,1} \eta_1\). In this case, (5.5.8) reduces to

\[(5.5.12)\quad ab = e + a_{1,0} \epsilon_1 + b_{0,1} \eta_1 + a_{1,0} b_{0,1} \epsilon_1 \eta_1.\]

Similarly, (5.5.9) and (5.5.10) reduce to

\[(5.5.13)\quad a^{-1} = e - a_{1,0} \epsilon_1, \quad b^{-1} = e - b_{0,1} \eta_1.\]

Thus

\[(5.5.14)\quad a^{-1} b^{-1} = e - a_{1,0} \epsilon_1 - b_{0,1} \eta_1 + a_{1,0} b_{0,1} \epsilon_1 \eta_1.\]

Combining (5.5.12) and (5.5.14), one can verify that

\[(5.5.15)\quad aba^{-1} b^{-1} = e + a_{1,0} b_{0,1} \epsilon_1 \eta_1 - b_{0,1} a_{1,0} \epsilon_1 \eta_1.\]

5.6 Differentiation

Let \(k\) be a commutative ring with a multiplicative identity element, and let \(T\) be an indeterminate. Also let \(A\) be a module over \(k\), so that the spaces \(A[T]\) and \(A[[T]]\) of formal polynomials and power series in \(T\) with coefficients in \(A\) can be defined as in Section 4.3. If \(f(T) = \sum_{j=0}^{\infty} f_j T^j \in A[[T]]\), then the formal derivative of \(f(T)\) is defined by

\[(5.6.1)\quad f'(T) = \sum_{j=1}^{\infty} j \cdot f_j T^{j-1} = \sum_{j=0}^{\infty} (j + 1) \cdot f_{j+1} T^j,\]

where \(j \cdot a\) is the sum of \(j\) \(a\)'s in \(A\) for every \(j \in \mathbb{Z}_+\) and \(a \in A\). Thus \(f'(T) \in A[[T]]\) too, and

\[(5.6.2)\quad f(T) \mapsto f'(T)\]

is a homomorphism from \(A[[T]]\) into itself, as a module over \(k\). If \(f(T) \in A[T]\), then \(f'(T) \in A[T]\) as well.

Let \(B\) and \(C\) be two more modules over \(k\), and let \(\beta\) be a mapping from \(A \times B\) into \(C\) that is bilinear over \(k\). Let \(f(T) \in A[[T]]\) be given as before, as well as \(g(T) = \sum_{i=0}^{\infty} g_i T^i \in B[[T]]\). Put

\[(5.6.3)\quad h_n = \sum_{j=0}^{n} \beta(f_j, g_{n-j}).\]
for every nonnegative integer \( n \), and

\[
\beta(f(T), g(T)) = h(T) = \sum_{n=0}^{\infty} h_n T^n,
\]

as in Section 4.5. We would like to verify that

\[
h'(T) = \beta(f'(T), g(T)) + \beta(f(T), g'(T)).
\]

This is the same as saying that

\[
(n + 1) \cdot h_{n+1} = \sum_{j=0}^{n+1} (n + 1) \cdot \beta(f_j, g_{n+1-j})
\]

\[
= \sum_{j=0}^{n+1} j \cdot \beta(f_j, g_{n+1-j}) + \sum_{j=0}^{n} (n + 1 - j) \cdot \beta(f_j, g_{n+1-j})
\]

\[
= \sum_{j=1}^{n+1} j \cdot \beta(f_j, g_{n+1-j}) + \sum_{j=0}^{n} (n + 1 - j) \cdot \beta(f_j, g_{n+1-j})
\]

\[
= \sum_{j=0}^{n} (j + 1) \cdot \beta(f_{j+1}, g_{n-j}) + \sum_{j=0}^{n} (n - j + 1) \cdot \beta(f_j, g_{n-j+1})
\]

for every \( n \geq 0 \). By the definition (5.6.3) of \( h_n \), we have that

\[
(n + 1) \cdot h_{n+1} = \sum_{j=0}^{n+1} (n + 1) \cdot \beta(f_j, g_{n+1-j})
\]

\[
= \sum_{j=0}^{n+1} j \cdot \beta(f_j, g_{n+1-j}) + \sum_{j=0}^{n} (n + 1 - j) \cdot \beta(f_j, g_{n+1-j})
\]

\[
= \sum_{j=1}^{n+1} j \cdot \beta(f_j, g_{n+1-j}) + \sum_{j=0}^{n} (n + 1 - j) \cdot \beta(f_j, g_{n+1-j})
\]

\[
= \sum_{j=0}^{n} (j + 1) \cdot \beta(f_{j+1}, g_{n-j}) + \sum_{j=0}^{n} (n - j + 1) \cdot \beta(f_j, g_{n-j+1})
\]

for each \( n \geq 0 \). This implies (5.6.6), as desired.

Let \( A \) be an algebra over \( k \) in the strict sense, where multiplication of \( a, b \in A \) is expressed as \( a \cdot b \). As in Section 4.6, multiplication on \( A \) can be extended to \( A[[T]] \), so that \( A[[T]] \) becomes an algebra over \( k \) in the strict sense too. If \( f(T) \) and \( g(T) \) are elements of \( A[[T]] \) and \( h(T) = f(T) g(T) \), then

\[
h'(T) = f'(T) g(T) + f(T) g'(T),
\]

as in (5.6.5). Thus (5.6.2) defines a derivation on \( A[[T]] \), as an algebra over \( k \). Similarly, the restriction of (5.6.2) to \( A[T] \) defines a derivation on \( A[T] \), as an algebra over \( k \).

Let \( A \) be a module over \( k \) again, and remember that \( A[[T]] \) may be considered as a module over \( k[[T]] \), as in Section 4.6. Let \( f(T) \in k[[T]] \) and \( g(T) \in A[[T]] \) be given, so that \( h(T) = f(T) g(T) \) defines an element of \( A[[T]] \) as well. Under these conditions, \( f'(T) \in k[[T]] \), \( g'(T) \), \( h'(T) \in A[[T]] \), and (5.6.8) holds in \( A[[T]] \). This may be considered as another instance of (5.6.5). More precisely, this uses scalar multiplication on \( A \) as a bilinear mapping from \( k \times A \) into \( A \), and its extension to \( k[[T]] \times A[[T]] \).
5.7 Polynomial functions

Let \( k \) be a commutative ring with multiplicative identity element, let \( A \) be a module over \( k \), and let \( T \) be an indeterminate. Also let \( f(T) = \sum_{j=0}^{n} f_j T^j \) be a formal polynomial in \( T \) with coefficients in \( A \), as in Section 4.3. If \( t \in k \), then

\[
(5.7.1) \quad f(t) = \sum_{j=0}^{n} f_j t^j
\]
defines an element of \( A \), where \( f_j t^j \) is defined using scalar multiplication on \( A \), and \( t^j \) is interpreted as being the multiplicative identity element 1 in \( k \) when \( j = 0 \). The mapping

\[
(5.7.2) \quad f(T) \mapsto f(t)
\]
defines a homomorphism from \( A[T] \) into \( A \), as modules over \( k \).

Let \( B \) and \( C \) be modules over \( k \) as well, and let \( \beta \) be a mapping from \( A \times B \) into \( C \) that is bilinear over \( k \). If \( f(T) \in A[T] \) and \( g(T) \in B[T] \), then

\[
(5.7.3) \quad h(T) = \beta(f(T), g(T))
\]
can be defined as an element of \( C[T] \) as in Section 4.5. Under these conditions, one can check that

\[
(5.7.4) \quad h(t) = \beta(f(t), g(t))
\]
for every \( t \in k \).

Now let \( A \) be an algebra over \( k \) in the strict sense, where multiplication of \( a, b \in A \) is expressed as \( a \cdot b \). Remember that multiplication on \( A \) can be extended to \( A[T] \), so that \( A[T] \) becomes an algebra over \( k \) in the strict sense, as in Section 4.6. If \( f(T), g(T) \in A[T] \) and \( h(T) = f(T) \cdot g(T) \), then

\[
(5.7.5) \quad h(t) = f(t) \cdot g(t)
\]
for every \( t \in k \), as in (5.7.4). Thus (5.7.2) defines a homomorphism from \( A[T] \) into \( A \), as algebras over \( k \).

Let \( A \) be a module over \( k \) again, so that \( A[T] \) may be considered as a module over \( k[T] \), as in Section 4.6. Let \( f(T) \in k[T] \) and \( g(T) \in A[T] \) be given, and let \( h(T) = f(T) \cdot g(T) \) be their product in \( A[T] \). If \( t \in k \), then \( f(t) \) is defined as an element of \( k \), \( g(t) \) and \( h(t) \) are defined as elements of \( A \), and (5.7.5) holds, as in (5.7.4).

Let \( t \in k \) be given, and suppose that \( \epsilon \in k \) satisfies

\[
(5.7.6) \quad \epsilon^2 = 0.
\]
Noe that

\[
(5.7.7) \quad (t + \epsilon)^j = t^j + j \cdot t^{j-1} \cdot \epsilon
\]
for every positive integer \( j \). If \( f(T) = \sum_{j=0}^{n} f_j T^j \in A[T] \), then

\[
(5.7.8) \quad f(t + \epsilon) = \sum_{j=0}^{n} f_j (t + \epsilon)^j
\]

\[
= \sum_{j=0}^{n} f_j t^j + \sum_{j=1}^{n} j \cdot f_j t^{j-1} \epsilon = f(t) + f'(t) \epsilon.
\]

Here \( f'(T) \in A[T] \) is as defined in the previous section, so that \( f'(t) \) is defined as an element of \( A \) as before.

Let \( A \) be an associative algebra over \( k \) with a multiplicative identity element \( e \). If \( f(T) = \sum_{j=0}^{n} f_j T^j \in k[T] \) and \( a \in A \), then

\[
(5.7.9) \quad f(a) = \sum_{j=0}^{n} f_j a^j
\]

is defined as an element of \( A \), where \( a^j \) is interpreted as being equal to \( e \) when \( j = 0 \). Let \( g(T) \) be another element of \( k[T] \), so that \( h(T) = f(T) g(T) \) is defined as an element of \( k[T] \) too. It is easy to see that

\[
(5.7.10) \quad h(a) = f(a) g(a),
\]

so that

\[
(5.7.11) \quad f(T) \mapsto f(a)
\]

defines a homomorphism from \( k[T] \) into \( A \), as algebras over \( k \).

Suppose that \( a, \epsilon \in A \) satisfy

\[
(5.7.12) \quad a \epsilon = \epsilon a
\]

and \( \epsilon^2 = 0 \). As in (5.7.7), we have that

\[
(5.7.13) \quad (a + \epsilon)^j = a^j + j \cdot a^{j-1} \epsilon
\]

for every \( j \in \mathbb{Z}_+ \). Using this, we get that

\[
(5.7.14) \quad f(a + \epsilon) = f(a) + f'(a) \epsilon,
\]

as in (5.7.8).

### 5.8 Several commuting indeterminates

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( A \) be a module over \( k \). Also let \( n \) be a positive integer, and let \( T_1, \ldots, T_n \) be \( n \) commuting indeterminates. As usual, a multi-index of length \( n \) is an \( n \)-tuple \( \alpha = (\alpha_1, \ldots, \alpha_n) \) of nonnegative integers, and we put

\[
(5.8.1) \quad |\alpha| = \alpha_1 + \cdots + \alpha_n.
\]
The corresponding formal monomial
\[ T^\alpha = T_1^{\alpha_1} \cdots T_n^{\alpha_n} \]
in \( T_1, \ldots, T_n \) has degree \(|\alpha|\). A formal power series in \( T_1, \ldots, T_n \) with coefficients in \( A \) can be expressed as
\[ f(T) = f(T_1, \ldots, T_n) = \sum_{\alpha \in (\mathbb{Z}_+ \cup \{0\})^n} f_\alpha T^\alpha, \]
where \( f_\alpha \in A \) for every multi-index \( \alpha \). The space \( A[[T_1, \ldots, T_n]] \) of all such formal power series can be defined as the space of all \( A \)-valued functions on the set \((\mathbb{Z}_+ \cup \{0\})^n\) of all multi-indices of length \( n \). This is a module over \( k \) with respect to pointwise addition and scalar multiplication of \( A \)-valued functions on \((\mathbb{Z}_+ \cup \{0\})^n\), which corresponds to termwise addition and scalar multiplication of formal power series as in (5.8.3). As a module over \( k \), \( A[[T_1, \ldots, T_n]] \) corresponds to the direct product of copies of \( A \) indexed by \((\mathbb{Z}_+ \cup \{0\})^n\).

A formal polynomial in \( T_1, \ldots, T_n \) with coefficients in \( A \) can be expressed as
\[ f(T) = f(T_1, \ldots, T_n) = \sum_{|\alpha| \leq N} f_\alpha T^\alpha, \]
where the sum is taken over multi-indices \( \alpha \) with \(|\alpha| \leq N \) for some nonnegative integer \( N \), and \( f_\alpha \in A \) for each such \( \alpha \). Of course, we can take \( f_\alpha = 0 \) when \(|\alpha| > N \), so that (5.8.4) may be considered as a formal power series in \( T_1, \ldots, T_n \) with coefficients in \( A \). The space \( A[T_1, \ldots, T_n] \) of all such formal polynomials can be defined as the space of all \( A \)-valued functions on \((\mathbb{Z}_+ \cup \{0\})^n\) that are equal to 0 for all but finitely many \( \alpha \), which is a submodule of \( A[[T_1, \ldots, T_n]] \).

As a module over \( k \), \( A[T_1, \ldots, T_n] \) corresponds to the direct sum of copies of \( A \) indexed by \((\mathbb{Z}_+ \cup \{0\})^n\). We can identify \( A \) with the submodule of \( A[T_1, \ldots, T_n] \) consisting of \( f(T) \) as in (5.8.4) with \( N = 0 \).

Suppose that \( A \) is an algebra over \( k \) in the strict sense, where multiplication of \( a, b \in A \) is expressed as \( ab \). Let \( f(T) \in A[[T_1, \ldots, T_n]] \) be as in (5.8.3), and let
\[ g(T) = \sum_{\beta \in (\mathbb{Z}_+ \cup \{0\})^n} g_\beta T^\beta \]
be another element of \( A[[T_1, \ldots, T_n]] \). If \( \alpha \) and \( \beta \) are multi-indices of length \( n \), then \( \alpha + \beta \) can be defined as a multi-index of length \( n \) by coordinatewise addition, and we have that
\[ |\alpha + \beta| = |\alpha| + |\beta|. \]
Put
\[ h_\gamma = \sum_{\alpha + \beta = \gamma} f_\alpha g_\beta \]
for every multi-index \( \gamma \) of length \( n \), where the sum is taken over all pairs of multi-indices \( \alpha, \beta \) such that \( \alpha + \beta = \gamma \). There are only finitely many such pairs,
so that the sum on the right side of (5.8.7) is a finite sum of elements of $A$.

Thus
\[ (5.8.8) \quad h(T) = \sum_{\gamma \in \mathbb{Z}_+ \cup \{0\}^n} h_\gamma T^\gamma \]
defines an element of $A[[T_1, \ldots, T_n]]$, and we put
\[ (5.8.9) \quad f(T) g(T) = h(T). \]

This extends multiplication on $A$ to $A[[T_1, \ldots, T_n]]$, so that the latter becomes an algebra over $k$ in the strict sense too. It is easy to see that $A[T_1, \ldots, T_n]$ is a subalgebra of $A[[T_1, \ldots, T_n]]$ with respect to this definition of multiplication. If multiplication on $A$ is commutative or associative, then one can check that multiplication on $A[[T_1, \ldots, T_n]]$ has the same property. If $A$ has a multiplicative identity element $e$, then the corresponding formal polynomial in $T_1, \ldots, T_n$ is the multiplicative identity element in $A[[T_1, \ldots, T_n]]$. In particular, $k[[T_1, \ldots, T_n]]$ is a commutative associative algebra over $k$.

Let $A$ be a module over $k$ again, let $f(T) \in k[[T_1, \ldots, T_n]]$ be as in (5.8.3), and let $g(T) \in A[[T_1, \ldots, T_n]]$ be as in (5.8.5). Thus $f_\alpha g_\beta$ is defined as an element of $A$ for all multi-indices $\alpha, \beta$, using scalar multiplication on $A$. If $\gamma$ is a multi-index of length $n$, then $h_\gamma$ can be defined as an element of $A$ as in (5.8.7). This permits us to define $h(T)$ as an element of $A[[T_1, \ldots, T_n]]$ as in (5.8.8), which can be used to define $f(T) g(T)$. One can verify that $A[[T_1, \ldots, T_n]]$ is a module over $k[[T_1, \ldots, T_n]]$ in this way. Similarly, if $f(T) \in k[T_1, \ldots, T_n]$ and $g(T) \in A[[T_1, \ldots, T_n]]$, then $h(T) \in A[[T_1, \ldots, T_n]]$. Using this definition of scalar multiplication, $A[T_1, \ldots, T_n]$ becomes a module over $k[T_1, \ldots, T_n]$.

Let $l$ and $m$ be positive integers, and let $X_1, \ldots, X_l, Y_1, \ldots, Y_m$ be commuting indeterminates. If $\beta$ and $\gamma$ are multi-indices of length $l$ and $m$, respectively, then let us identify $(\beta, \gamma)$ with a multi-index of length $l + m$. This corresponds to identifying $(\mathbb{Z}_+ \cup \{0\})^l \times (\mathbb{Z}_+ \cup \{0\})^m$ with $(\mathbb{Z}_+ \cup \{0\})^{l+m}$. We may consider
\[ (5.8.10) \quad X^\beta Y^\gamma = X_1^{\beta_1} \cdots X_l^{\beta_l} Y_1^{\gamma_1} \cdots Y_m^{\gamma_m} \]
as a formal monomial in the variables $X_1, \ldots, X_l, Y_1, \ldots, Y_m$ of degree $|\beta| + |\gamma|$. As before, $A[[X_1, \ldots, X_l]]$ is a module over $k$, so that
\[ (5.8.11) \quad A[[X_1, \ldots, X_l]][Y_1, \ldots, Y_m] \]
can be defined as a module over $k$ as well. There is a simple one-to-one correspondence between the elements of (5.8.11) and
\[ (5.8.12) \quad A[[X_1, \ldots, X_l, Y_1, \ldots, Y_m]], \]
which defines an isomorphism between these modules over $k$. This correspondence takes
\[ (5.8.13) \quad (A[X_1, \ldots, X_l])[Y_1, \ldots, Y_m] \]
on to
\[ (5.8.14) \quad A[X_1, \ldots, X_l, Y_1, \ldots, Y_m]. \]
If $A$ is an algebra over $k$ in the strict sense, then we get an isomorphism between (5.8.11) and (5.8.12) as algebras over $k$. In particular, we get an isomorphism between
\begin{equation}
(k[[X_1,\ldots,X_l]])[[Y_1,\ldots,Y_m]]
\end{equation}
and
\begin{equation}
k[[X_1,\ldots,X_l,Y_1,\ldots,Y_m]]
\end{equation}
as algebras over $k$. Similarly, if $A$ is a module over $k$, then scalar multiplication on (5.8.11) by elements of (5.8.15) corresponds to scalar multiplication on (5.8.12) by elements of (5.8.16).

### 5.9 Polynomial functions in several variables

Let $k$ be a commutative ring with a multiplicative identity element, and let $n$ be a positive integer. As usual, we let $k^n$ be the space of $n$-tuples of elements of $k$. If $t = (t_1,\ldots,t_n) \in k^n$ and $\alpha$ is a multi-index of length $n$, then $t^\alpha$ is defined as an element of $k$ by
\begin{equation}
t^\alpha = t_1^{\alpha_1} \cdots t_n^{\alpha_n}.
\end{equation}
Here $t_j^0$ is interpreted as being the multiplicative identity element $1$ in $k$ when $\alpha_j = 0$, as before. If $\beta$ is another multi-index of length $n$, then
\begin{equation}
t^{\alpha+\beta} = t^\alpha t^\beta.
\end{equation}

Let $A$ be a module over $k$, and let $T_1,\ldots,T_n$ be $n$ commuting indeterminates. Also let $f(T)$ be a formal polynomial in $T_1,\ldots,T_n$ with coefficients in $A$, as in (5.8.4). If $t \in k^n$, then
\begin{equation}
f(t) = \sum_{|\alpha| \leq N} f_\alpha t^\alpha
\end{equation}
defines an element of $A$, where $f_\alpha t^\alpha$ is defined using scalar multiplication on $A$ for each multi-index $\alpha$. The mapping
\begin{equation}
f(T) \mapsto f(t)
\end{equation}
defines a homomorphism from $A[T_1,\ldots,T_n]$ into $A$, as modules over $k$.

Let $A$ be an algebra over $k$ in the strict sense, where multiplication of $a, b \in A$ is expressed as $ab$. Remember that multiplication on $A$ can be extended to $A[T_1,\ldots,T_n]$, as in the previous section. Let $f(T), g(T) \in A[T_1,\ldots,T_n]$ be given, and put $h(T) = f(T) g(T)$. If $t \in k^n$, then one can check that
\begin{equation}
h(t) = f(t) g(t).
\end{equation}
This means that (5.9.4) defines a homomorphism from $A[T_1,\ldots,T_n]$ into $A$, as algebras over $k$.

Let $A$ be a module over $k$ again, and remember that $A[T_1,\ldots,T_n]$ may be considered as a module over $k[T_1,\ldots,T_n]$, as in the previous section. Let $f(T)$
114  CHAPTER 5. SOME RELATED NOTIONS

in \( k[T_1, \ldots, T_n] \) and \( g(T) \in A[T_1, \ldots, T_n] \) be given, so that \( h(T) = f(T) g(T) \) is defined as an element of \( A[T_1, \ldots, T_n] \) as well. If \( t \in k^n \), then \( f(t) \in k \), \( g(t) \) and \( h(t) \) are elements of \( A \), and one can verify that (5.9.5) holds.

Let \( A \) be an associative algebra over \( k \) with a multiplicative identity element \( e \), and let \( A^n \) be the space of \( n \)-tuples of elements of \( A \). Suppose that \( a = (a_1, \ldots, a_n) \in A^n \) has commuting coordinates, so that

\[
\begin{align*}
\alpha_j a_l &= a_l a_j
\end{align*}
\]

(5.9.6)

for all \( j, l = 1, \ldots, n \). Of course, this condition holds trivially when \( n = 1 \). If \( \alpha \) is a multi-index of length \( n \), then \( a^\alpha \) is defined as an element of \( A \) by

\[
\begin{align*}
a^\alpha &= a_1^{\alpha_1} \cdots a_n^{\alpha_n},
\end{align*}
\]

(5.9.7)

where \( a_j^{\alpha_j} \) is interpreted as being equal to \( e \) when \( \alpha_j = 0 \). Note that

\[
\begin{align*}
a^{\alpha + \beta} &= a^\alpha a^\beta
\end{align*}
\]

(5.9.8)

for all multi-indices \( \alpha, \beta \) under these conditions. Let \( f(T) \) be a formal polynomial in \( T_1, \ldots, T_n \) with coefficients in \( k \), as in (5.8.4) again. As before, \( f(a) \) is defined as an element of \( A \) by

\[
\begin{align*}
f(a) &= \sum_{|\alpha| \leq N} f_\alpha a^\alpha,
\end{align*}
\]

(5.9.9)

where \( f_\alpha a^\alpha \) is defined using scalar multiplication on \( A \). If \( g(T) \in k[T_1, \ldots, T_n] \) too and \( h(T) = f(T) g(T) \), then one can verify that

\[
\begin{align*}
h(a) &= f(a) g(a).
\end{align*}
\]

(5.9.10)

It follows that \( f(T) \mapsto f(a) \) is a homomorphism from \( k[T_1, \ldots, T_n] \) into \( A \), as algebras over \( k \), since this mapping is clearly linear over \( k \).

5.10  Partial derivatives

Let \( n \) be a positive integer, let \( \alpha \) be a multi-index of length \( n \), and let \( l \) be a positive integer with \( l \leq n \). The multi-index \( \alpha(l) \) of length \( n \) is defined by

\[
\begin{align*}
\alpha_j(l) &= \alpha_j \quad \text{when } j \neq l
\end{align*}
\]

(5.10.1)

and

\[
\begin{align*}
\alpha_l(l) &= \alpha_l - 1 \quad \text{when } \alpha_l \geq 1 \\
&= 0 \quad \text{when } \alpha_l = 0.
\end{align*}
\]

(5.10.2)

Similarly, let \( \alpha^+(l) \) be the multi-index of length \( n \) defined by

\[
\begin{align*}
\alpha_j^+(l) &= \alpha_j \quad \text{when } j \neq l \\
&= \alpha_l + 1 \quad \text{when } j = l.
\end{align*}
\]

(5.10.3)
5.10. PARTIAL DERIVATIVES

Let $k$ be a commutative ring with a multiplicative identity element, let $A$ be a module over $k$, and let $T_1, \ldots, T_n$ be $n$ commuting indeterminates. Also let $f(T)$ be a formal power series in $T_1, \ldots, T_n$ with coefficients in $A$, as in (5.8.3). The formal partial derivative of $f(T)$ in $T_l$ can be defined as a formal power series in $T_1, \ldots, T_n$ with coefficients in $A$ by

$$\frac{\partial f(T)}{\partial T_l} = \sum_{\alpha \in (\mathbb{Z}_+ \cup \{0\})^n} (\alpha_l + 1) \cdot f_{\alpha(l)} T^\alpha. \quad (5.10.4)$$

This is basically the same as

$$\sum_{\alpha \in (\mathbb{Z}_+ \cup \{0\})^n} \alpha_l \cdot f_{\alpha(l)} = \sum_{\alpha_l \geq 1} \alpha_l \cdot f_{\alpha(l)}, \quad (5.10.5)$$

where the second sum is taken over all multi-indices $\alpha$ with $\alpha_l \geq 1$. Note that

$$f(T) \mapsto \partial_l f(T) \quad (5.10.6)$$

defines a homomorphism from $A[[T_1, \ldots, T_n]]$ into itself, as a module over $k$. Of course, if $f(T) \in A[T_1, \ldots, T_n]$, then the previous sums reduce to finite sums, and $\partial_l f(T) \in A[T_1, \ldots, T_n]$. One can check that

$$\partial_l (\partial_m f(T)) = \partial_m (\partial_l f(T)) \quad (5.10.7)$$

for every $l, m = 1, \ldots, n$ and $f(T) \in A[[T_1, \ldots, T_n]]$.

If $n = 1$, then (5.10.4) reduces to the definition of the derivative in Section 5.6. If $n > 1$, then we can identify $f(T) \in A[[T_1, \ldots, T_n]]$ with a formal power series in $T_l$ whose coefficients are formal power series in the other variables $T_j, j \neq l$, with coefficients in $A$, as in Section 5.8. The derivative of this formal power series in $T_l$ can be defined as in Section 5.6, as a formal power series in $T_l$ whose coefficients are formal power series in the other variables. This differentiated formal power series can be identified with a formal power series in $T_1, \ldots, T_n$, as before, which is the same as (5.10.4).

Suppose that $A$ is an algebra over $k$ in the strict sense, so that $A[[T_1, \ldots, T_n]]$ is an algebra over $k$ in the strict sense as well, as in Section 5.8. Under these conditions, (5.10.6) defines a derivation on $A[[T_1, \ldots, T_n]]$. This can be reduced to the analogous statement for polynomials in one variable in Section 5.6, as in the preceding paragraph, or verified directly as in the $n = 1$ case.

Let $A$ be a module over $k$ again, and remember that $A[[T_1, \ldots, T_n]]$ may be considered as a module over $k[[T_1, \ldots, T_n]]$, as in Section 5.8. If $f(T)$ is an element of $k[[T_1, \ldots, T_n]]$ and $g(T) \in A[[T_1, \ldots, T_n]]$, then

$$\partial_l (f(T) g(T)) = (\partial_l f(T)) g(T) + f(T) (\partial_l g(T)), \quad (5.10.8)$$

as elements of $A[[T_1, \ldots, T_n]]$. This can be reduced to the analogous statement for polynomials in one variable in Section 5.6, as before, or verified directly in a similar way.
Let \( t \in k^n \) be given, and suppose that \( u \in k^n \) satisfies

\[
(5.10.9) \quad u_j \, u_l = 0
\]

for all \( j, l = 1, \ldots, n \). Of course, \( t + u \) is defined as an element of \( k^n \), using coordinate-wise addition. Let \( \alpha \) be a multi-index of length \( n \), so that \( t^\alpha \) and \((t + u)^\alpha\) are defined as elements of \( k \), as in (5.9.1). As in (5.7.7),

\[
(5.10.10) \quad (t_l + u_l)^\alpha_l = t_l^{\alpha_l} + \alpha_l \cdot t_l^{\alpha_l} + u_l = (t_l + u_l)^\alpha_l
\]

for each \( l = 1, \ldots, n \) when \( \alpha_l \neq 0 \), because \( u_l^2 = 0 \). If \( \alpha_l = 0 \), then \((t_l + u_l)^\alpha_l = t_l^\alpha_l \). Thus

\[
(5.10.11) \quad (t_l + u_l)^\alpha_l = t_l^{\alpha_l} + \alpha_l \cdot t_l^{\alpha_l} + u_l
\]

for every \( l = 1, \ldots, n \). Using this, one can check that

\[
(5.10.12) \quad (t + u)^\alpha = t^\alpha + \sum_{l=1}^n \alpha_l \cdot t^{\alpha_l} \cdot u_l
\]

If \( f(T) \in A[T_1, \ldots, T_n] \), then it follows that

\[
(5.10.13) \quad f(t + u) = f(t) + \sum_{l=1}^n \left( \partial_l f \right)(t) \cdot u_l
\]

Let \( A \) be an associative algebra over \( k \) with a multiplicative identity element \( e \), and suppose that \( a \in A^n \) has commuting coordinates, as in the previous section. Also let \( u \) be an element of \( A^n \) that satisfies (5.10.9), and whose coordinates commute with the coordinates of \( a \), so that

\[
(5.10.14) \quad a_j \, u_l = u_l \, a_j
\]

for all \( j, l = 1, \ldots, n \). Note that \( u \) has commuting coordinates, and hence \( a + u \) has commuting coordinates. Let \( \alpha \) be a multi-index of length \( n \) again, so that \( a^\alpha \) and \((a + u)^\alpha\) are defined as elements of \( A \) as before. As in (5.10.11),

\[
(5.10.15) \quad (a_l + u_l)^\alpha_l = a_l^{\alpha_l} + \alpha_l \cdot a_l^{\alpha_l} + u_l
\]

for every \( l = 1, \ldots, n \). This implies that

\[
(5.10.16) \quad (a + u)^\alpha = a^\alpha + \sum_{l=1}^n \alpha_l \cdot a^{\alpha_l} \cdot u_l,
\]

as before. If \( f(T) \in k[T_1, \ldots, T_n] \), then we get that

\[
(5.10.17) \quad f(a + u) = f(a) + \sum_{l=1}^n \left( \partial_l f \right)(a) \cdot u_l.
\]
5.11 Formal differential operators

Let $k$ be a commutative ring with a multiplicative identity element, and let $n$ be a positive integer. Also let $\partial_1, \ldots, \partial_n$ be commuting formal symbols, which may be used to represent partial derivatives, as in the previous section. If $\alpha$ is a multi-index of length $n$, then let

$$\partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}$$

be the corresponding formal product of $\partial_i$’s.

Let $T_1, \ldots, T_n$ be $n$ commuting indeterminates, as before. A formal differential operator in $\partial_1, \ldots, \partial_n$ with coefficients in $k[[T_1, \ldots, T_n]]$ can be expressed as

$$\sum_{|\alpha| \leq N} a^\alpha(T) \partial^\alpha,$$

where $N$ is a nonnegative integer, the sum is taken over all multi-indices $\alpha$ of length $n$ with $|\alpha| \leq N$, and $a^\alpha(T) \in k[[T_1, \ldots, T_n]]$ for each such $\alpha$. As usual, we can take $a^\alpha(T) = 0$ when $|\alpha| > N$, so that $a^\alpha(T)$ is defined for every multi-index $\alpha$. This is a module over $k$ with respect to pointwise addition and scalar multiplication, which corresponds to termwise addition and scalar multiplication of sums as in (5.11.2). As a module over $k$, this corresponds to the direct sum of copies of $k[[T_1, \ldots, T_n]]$ with the space of formal differential operators as a subalgebra.

Multiplication on $k[[T_1, \ldots, T_n]]$ can be extended to these formal differential operators, with

$$\partial_l(b^\beta(T) \partial^\beta) = (\partial_l b^\beta(T)) \partial^\beta + b^\beta(T) \partial_l \partial^\beta$$

for every $l = 1, \ldots, n$, multi-index $\beta$, and $b^\beta(T) \in k[[T_1, \ldots, T_n]]$. Note that

$$\partial_l \partial^\beta = \partial^{\beta + (l)},$$

in the notation of the previous section. The space of these formal differential operators is an associative algebra over $k$ in this way, which contains $k[[T_1, \ldots, T_n]]$ as a subalgebra. The multiplicative identity element of $k$ is also the multiplicative identity element in the space of these formal differential operators, when considered as an element of $k[[T_1, \ldots, T_n]]$ and hence a formal differential operator, as before.

A formal differential operator in $\partial_1, \ldots, \partial_n$ with coefficients in $k[T_1, \ldots, T_n]$ can be expressed as in (5.11.2), with $a^\alpha(T) \in k[T_1, \ldots, T_n]$ for each $\alpha$. The space of these formal differential operators is a subalgebra of the space of formal differential operators in $\partial_1, \ldots, \partial_n$ with coefficients in $k[[T_1, \ldots, T_n]]$. The space of formal differential operators in $\partial_1, \ldots, \partial_n$ with coefficients in $T_1, \ldots, T_n$ contains $k[T_1, \ldots, T_n]$ as a subalgebra, as before.
CHAPTER 5. SOME RELATED NOTIONS

Let $A$ be a module over $k$. If $\alpha$ is a multi-index of length $n$ and $f(T)$ is a formal power series in $T_1, \ldots, T_n$ with coefficients in $A$, then

\[ \partial^\alpha f(T) = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f(T) \]

(5.11.5)

defines an element of $A[[T_1, \ldots, T_n]]$ as well, where partial derivatives are defined on $A[[T_1, \ldots, T_n]]$ as in the previous section. This is interpreted as being equal to $f(T)$ when $\alpha = 0$. Similarly, if (5.11.2) is a formal differential operator in $\partial_1, \ldots, \partial_n$ with coefficients in $k[[T_1, \ldots, T_n]]$, then

\[ \left( \sum_{|\alpha| \leq N} a^\alpha(T) \partial^\alpha \right) f(T) = \sum_{|\alpha| \leq N} a^\alpha(T) \partial^\alpha f(T) \]

(5.11.6)

defines an element of $A[[T_1, \ldots, T_n]]$. Thus (5.11.2) induces a mapping from $A[[T_1, \ldots, T_n]]$ into itself, which is linear over $k$. If the coefficients $a^\alpha(T)$ of (5.11.2) are elements of $k[T_1, \ldots, T_n]$ and $f(T) \in A[[T_1, \ldots, T_n]]$, then (5.11.6) is in $A[T_1, \ldots, T_n]$ too.

Remember that the space

\[ \text{Hom}_k(A[[T_1, \ldots, T_n]], A[[T_1, \ldots, T_n]]) \]

(5.11.7)

of homomorphisms from $A[[T_1, \ldots, T_n]]$ into itself, as a module over $k$, is an algebra over $k$ with respect to composition of mappings. The remarks in the preceding paragraph define a mapping from the space of formal differential operators in $\partial_1, \ldots, \partial_n$ with coefficients in $k[[T_1, \ldots, T_n]]$ into (5.11.7). One can check that this mapping is an algebra homomorphism, with respect to multiplication of formal differential operators, as mentioned earlier.

Suppose now that $A = k$, as a module over itself. In this case, one can verify that a formal differential operator (5.11.2) in $\partial_1, \ldots, \partial_n$ with coefficients in $k[[T_1, \ldots, T_n]]$ is uniquely determined by the corresponding mapping from $k[[T_1, \ldots, T_n]]$ into itself. More precisely, (5.11.2) is uniquely determined by the restriction of this mapping to $k[T_1, \ldots, T_n]$.

5.12 First-order differential operators

Let $k$ be a commutative ring with a multiplicative identity element, let $n$ be a positive integer, and let $\partial_1, \ldots, \partial_n$ be commuting formal symbols, as in the previous section. Also let $T_1, \ldots, T_n$ be commuting indeterminates, and let

\[ a(T) = (a^1(T), \ldots, a^n(T)) \]

(5.12.1)

be an $n$-tuple of formal power series in $T_1, \ldots, T_n$ with coefficients in $k$. Put

\[ D_{a(T)} = \sum_{j=1}^n a^j(T) \partial_j \]

(5.12.2)

which defines a formal partial differential operator in $\partial_1, \ldots, \partial_n$ with coefficients in $k[[T_1, \ldots, T_n]]$, as in the previous section.
Let \( b(T) = (b^1(T), \ldots, b^n(T)) \) be another \( n \)-tuple of formal power series in \( T_1, \ldots, T_n \) with coefficients in \( k \), so that \( D_{b(T)} \) can be defined as before. The products \( D_{a(T)} D_{b(T)} \) and \( D_{b(T)} D_{a(T)} \) can be defined as formal differential operators in \( \partial_1, \ldots, \partial_n \) with coefficients in \( k[[T_1, \ldots, T_n]] \) as well, as in the previous section. It is easy to see that

\[
D_{a(T)} D_{b(T)} - D_{b(T)} D_{a(T)} = \sum_{j=1}^n \sum_{l=1}^n (a^j(T) \partial_j b^l(T) - b^l(T) \partial_j a^j(T)) \partial_l,
\]

using (5.11.3). Put \( c(T) = (c^1(T), \ldots, c^n(T)) \), where

\[
c^l(T) = \sum_{j=1}^n (a^j(T) \partial_j b^l(T) - b^l(T) \partial_j a^j(T))
\]

for each \( l = 1, \ldots, n \). Thus \( c(T) \) is another \( n \)-tuple of elements of \( k[[T_1, \ldots, T_n]] \), and

\[
D_{a(T)} D_{b(T)} - D_{b(T)} D_{a(T)} = D_{c(T)}.
\]

If \( f(T) \in k[[T_1, \ldots, T_n]] \), then

\[
D_{a(T)} f(T) = \sum_{j=1}^n a^j(T) \partial_j f(T)
\]

defines an element of \( k[[T_1, \ldots, T_n]] \) too, as in the previous section. One can check that this defines a derivation on \( k[[T_1, \ldots, T_n]] \), because partial derivatives define derivations on \( k[[T_1, \ldots, T_n]] \).

Let \( \delta \) be any derivation on \( k[[T_1, \ldots, T_n]] \), as an algebra over \( k \). Note that \( T_j \) may be considered as an element of \( k[[T_1, \ldots, T_n]] \) for each \( j = 1, \ldots, n \), where more precisely the coefficient of \( T_j \) is the multiplicative identity element in \( k \). Thus

\[
a^j(T) = \delta(T_j)
\]

defines an element of \( k[T_1, \ldots, T_n] \) for each \( j = 1, \ldots, n \). This permits us to define \( a(T) \) as an \( n \)-tuple of elements of \( k[T_1, \ldots, T_n] \) as in (5.12.1), so that \( D_{a(T)} \) can be defined as in (5.12.2). If \( f(T) \in k[T_1, \ldots, T_n] \), then one can verify that

\[
\delta(f(T)) = D_{a(T)} f(T),
\]

where the right side is defined as in (5.12.6).

Now let \( \delta \) be a derivation on \( k[[T_1, \ldots, T_n]] \), as an algebra over \( k \). In this case, (5.12.7) defines an element of \( k[[T_1, \ldots, T_n]] \) for each \( j = 1, \ldots, n \), so that we can define \( a(T) \) as an \( n \)-tuple of elements of \( k[[T_1, \ldots, T_n]] \) as in (5.12.1). Thus \( D_{a(T)} \) can be defined as in (5.12.2), and (5.12.8) holds for every \( f(T) \in k[T_1, \ldots, T_n] \), as in the preceding paragraph. Of course, we would like to extend this to \( f(T) \in k[[T_1, \ldots, T_n]] \).
Let us say that $f(T) \in \sum_{\alpha \in \mathbb{Z}_{+ \cup \{0\}}} f_{\alpha} T^\alpha \in k[[T_1, \ldots, T_n]]$ vanishes to order $L$ for some nonnegative integer $L$ if $f_\alpha = 0$ for every multi-index $\alpha$ with $|\alpha| \leq L$. If $g(T) \in k[[T_1, \ldots, T_n]]$, then
\begin{equation}
 g(T) T^\beta
\end{equation}
vanishes to order $|\beta| - 1$ for every nonzero multi-index $\beta$. If $f(T)$ vanishes to order $L$ for some $L \geq 0$, then $f(T)$ can be expressed as a finite sum of elements of $k[[T_1, \ldots, T_n]]$ of the form (5.12.9), when $|\beta| = L + 1$.

If $f(T)$ vanishes to order $L$ for some $L \geq 1$, then $\delta(f(T))$ vanishes to order $L - 1$. This can be verified directly when $f(T)$ is of the form (5.12.9) with $|\beta| = L + 1$, and otherwise one can reduce to that case, as in the preceding paragraph. One can use this to get that (5.12.8) holds for every $f(T) \in k[[T_1, \ldots, T_n]]$, as desired.

### 5.13 Homogeneous formal polynomials

Let $k$ be a commutative ring with a multiplicative identity element, let $T_1, \ldots, T_n$ be $n$ commuting indeterminates for some positive integer $n$, and let $A$ be a module over $k$. A formal polynomial $f(T)$ in $T_1, \ldots, T_n$ with coefficients in $A$ is said to be homogeneous of degree $d$ for some nonnegative integer $d$ if $f(T)$ can be expressed as
\begin{equation}
 f(T) = \sum_{|\alpha| = d} f_{\alpha} T^\alpha,
\end{equation}
where the sum is taken over all multi-indices $\alpha$ of length $n$ with $|\alpha| = d$, and $f_\alpha$ is an element of $A$ for all such $\alpha$. Equivalently, this means that the coefficient $f_\alpha$ of $T^\alpha$ in $f(T)$ is equal to 0 when $|\alpha| \neq d$. The space $A_d[T_1, \ldots, T_n]$ of these formal polynomials is a submodule of $A[T_1, \ldots, T_n]$, as a module over $k$. Note that $A_d[T_1, \ldots, T_n]$ corresponds to a direct sum of copies of $A$ indexed by multi-indices $\alpha$ with $|\alpha| = d$, and that $A[T_1, \ldots, T_n]$ can be viewed as the direct sum of $A_d[T_1, \ldots, T_n]$ over all nonnegative integers $d$. Similarly, $A[[T_1, \ldots, T_n]]$ can be viewed as the direct product of $A_d[T_1, \ldots, T_n]$ over all nonnegative integers $d$. If $f(T) \in A_d[T_1, \ldots, T_n]$ for some $d \geq 1$, then it is easy to see that $\partial_l f(T)$ is homogeneous of degree $d - 1$ for every $l = 1, \ldots, n$.

Suppose that $A$ is an algebra over $k$ in the strict sense. If $f(T)$, $g(T)$ are homogeneous formal polynomials in $T_1, \ldots, T_n$ with coefficients in $A$ for some nonnegative integers $d_1$, $d_2$, respectively, then one can check that $f(T)g(T)$ is homogeneous of degree $d_1 + d_2$.

Let us now take $A = k$, and let
\begin{equation}
 a^j(T) = \sum_{l=1}^n a^j_l T_l
\end{equation}
be elements of $k_1[T_1, \ldots, T_n]$ for $j = 1, \ldots, n$, so that $a^j_l \in k$ for all $j, l = 1, \ldots, n$. Thus $\partial_i a^j_l(T) = a^j_l$ for every $j, l = 1, \ldots, n$. Put $a(T) = (a^1(T), \ldots, a^n(T))$, and
5.14. HOMOGENEOUS DIFFERENTIAL OPERATORS

Let \( \partial_1, \ldots, \partial_n \) be commuting formal symbols, as in the previous two sections. Consider the corresponding formal differential operator \( D_{a(T)} \) in \( \partial_1, \ldots, \partial_n \), as before. If \( f(T) \) is a formal polynomial in \( T_1, \ldots, T_n \) with coefficients in \( k \), then \( D_{a(T)} f(T) \) is defined as an element of \( k[T_1, \ldots, T_n] \) too. More precisely, if \( f(T) \) is homogeneous of degree \( d \), then \( D_{a(T)} f(T) \) is homogeneous of degree \( d \) as well.

In particular, if \( f(T) = \sum_{j=1}^{n} f_j T_j \) is homogeneous of degree 1, so that \( f_j \in k \) for \( j = 1, \ldots, n \), then

\[
D_{a(T)} f(T) = \sum_{j=1}^{n} a^j(T) \partial_j f(T) = \sum_{j=1}^{n} \sum_{l=1}^{n} a^l_j f_j T_l.
\]

Let

\[
b^j(T) = \sum_{l=1}^{n} b^l_j T_l
\]

be elements of \( k[T_1, \ldots, T_n] \) for \( j = 1, \ldots, n \), and put \( b(T) = (b^1(T), \ldots, b^n(T)) \).

Put

\[
c^j(T) = \sum_{l=1}^{n} (a^j(T) b^l_l - b^l(T) a^j_l) = \sum_{l=1}^{n} \sum_{m=1}^{n} (a^l_m b^j_l - b^l_m a^j_l) T_m
\]

for each \( j = 1, \ldots, n \), and \( c(T) = (c^1(T), \ldots, c^n(T)) \). Note that \( c^j(T) \) is an element of \( k[T_1, \ldots, T_n] \) for \( j = 1, \ldots, n \). If \( D_{b(T)} \) and \( D_{c(T)} \) are the formal differential operators in \( \partial_1, \ldots, \partial_n \) corresponding to \( b(T) \) and \( c(T) \), respectively, then

\[
D_{a(T)} D_{b(T)} - D_{b(T)} D_{a(T)} = D_{c(T)},
\]

as in (5.12.5).

5.14 Homogeneous differential operators

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( n \) be a positive integer. Also let \( T_1, \ldots, T_n \) be commuting indeterminates, and let \( \partial_1, \ldots, \partial_n \) be commuting formal symbols, as in Section 5.11. Consider a formal differential operator

\[
L = \sum_{|\alpha|\leq N} a_\alpha(T) \partial^\alpha
\]

in \( \partial_1, \ldots, \partial_n \) with coefficients in \( k[T_1, \ldots, T_n] \). Let us say that \( L \) is homogeneous of degree \( d \) for some integer \( d \) if the following conditions hold. If \( \alpha \) is a multi-index such that \( d \geq |\alpha| \), then \( a_\alpha(T) \) should be homogeneous of degree \( d + |\alpha| \) as a formal polynomial in \( T_1, \ldots, T_n \) with coefficients in \( k \). Otherwise, if \( d \) is strictly less than \( |\alpha| \), then \( a_\alpha(T) = 0 \). The space of these formal differential operators that are homogeneous of degree \( d \) is a submodule of the space of all
formal differential operators in $\partial_1, \ldots, \partial_n$ with coefficients in $k[T_1, \ldots, T_n]$, as a module over $k$.

Let $L_1$ and $L_2$ be formal differential operators in $\partial_1, \ldots, \partial_n$ with coefficients in $k[T_1, \ldots, T_n]$, so that their product $L_1 L_2$ can be defined as a formal differential operator with coefficients in $k[T_1, \ldots, T_n]$ as well. If $L_1, L_2$ are homogeneous of degrees $d_1, d_2 \in \mathbb{Z}$, respectively, then one can check that $L_1 L_2$ is homogeneous of degree $d_1 + d_2$. More precisely, one can start with the case where $L_1 = \partial_j$ for some $j = 1, \ldots, n$, so that $d_1 = -1$. Using this, one can obtain the analogous statement for $L_1 = \partial^\alpha$ for some multi-index $\alpha$, so that $d_1 = -|\alpha|$. One can use this to obtain the analogous statement when $L_1$ is homogeneous of any degree $d_1$.

Let $L_1$ be a formal differential operator in $\partial_1, \ldots, \partial_n$ with coefficients in $k[T_1, \ldots, T_n]$ again, and let $A$ be a module over $k$. Also let $f(T)$ be a formal polynomial in $T_1, \ldots, T_n$ with coefficients in $A$, so that $L_1 f(T)$ is defined as an element of $A[T_1, \ldots, T_n]$ too. Suppose that $L_1$ is homogeneous of degree $d_1 \in \mathbb{Z}$, and that $f(T)$ is homogeneous of degree $d$ for some nonnegative integer $d$. One can verify that $L_1 f(T)$ is homogeneous of degree $d_1 + d$ when $d_1 \geq -d$, and that $L_1 f(T) = 0$ otherwise. Indeed, if $L_1 = \partial_j$ for some $j = 1, \ldots, n$, then this was mentioned in the previous section. As before, one can use this to obtain the analogous statement when $L_1 = \partial^\alpha$ for some multi-index $\alpha$. The analogous statement for any $L_1$ follows easily from this.

Remember that the space of formal differential operators in $\partial_1, \ldots, \partial_n$ with coefficients in $k[T_1, \ldots, T_n]$ is an algebra over $k$. The collection of such formal differential operators that are homogeneous of degree 0 is a subalgebra of this algebra. This subalgebra is generated as an an algebra over $k$ by homogeneous differential operators of degree 0 as in (5.14.1) with $N = 1$. To see this, one can start with a homogeneous differential operator of degree 0 of the form $T^\alpha \partial^\beta$, where $\alpha, \beta$ are multi-indices with $|\alpha| = |\beta| \geq 2$. This can be approximated by a product of $|\alpha| = |\beta|$ operators of the form $T_j \partial_l$ for some $j, l = 1, \ldots, n$. More precisely, one can choose the approximation so that the difference is a formal differential operator of lower order. One can repeat the process to express any homogeneous differential operator of degree 0 as a finite sum of products of homogeneous differential operators of degree 0 and order less than or equal to 1, as desired.
Chapter 6

Bilinear actions and representations

6.1 Bilinear actions

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$, $V$ be modules over $k$. Also let $\beta$ be a mapping from $A \times V$ into $V$ that is bilinear over $k$. This may be described as a bilinear action of $A$ on $V$ over $k$, or more precisely as a bilinear action of $A$ on $V$ on the left. It is sometimes convenient to consider a mapping from $V \times A$ into $V$ that is bilinear over $k$ as a bilinear action of $A$ on $V$ on the right. If $A$ is an associative algebra over $k$, or a Lie algebra over $k$, then we may be interested in bilinear actions that satisfy additional conditions, as in Sections 6.4 and 6.5.

Alternatively, we may use the notation

(6.1.1) $\rho_a(v) = \beta(a, v),$

where $a \in A$ and $v \in V$. The linearity of $\beta$ means that $\rho_a$ is a module homomorphism from $V$ into itself for each $a \in A$, and that $a \mapsto \rho_a$ defines a module homomorphism from $A$ into the space $\text{Hom}_k(V, V)$ of module homomorphisms from $V$ into itself. We may use $\rho$ to denote a module homomorphism from $A$ into $\text{Hom}_k(V, V)$ in this way, which defines a bilinear action of $A$ on $V$ as in (6.1.1). We may also use the notation

(6.1.2) $a \cdot v = \beta(a, v)$

for $a \in A$ and $v \in V$. A bilinear action of $A$ on $V$ on the right may be expressed by $v \cdot a$ for $a \in A$ and $v \in V$.

Let $T$ be an indeterminate, and remember that $k[T]$, $A[T]$, and $V[T]$ are the corresponding spaces of formal polynomials in $T$ with coefficients in $k$, $A$, and $V$, respectively. A bilinear action of $A$ on $V$ over $k$ can be extended to a bilinear action of $A[T]$ on $V[T]$ over $k[T]$. More precisely, one could start with a
mapping from $A \times V$ into $V[T]$ that is bilinear over $k$. This can be extended to a mapping from $A[T] \times V[T]$ into $V[T]$ that is bilinear over $k[T]$, as in Sections 4.5 and 4.6.

Remember that $k[[T]]$, $A[[T]]$, and $V[[T]]$ are the corresponding spaces of formal power series in $T$ with coefficients in $k$, $A$, and $T$, respectively. A bilinear action of $A$ on $V$ over $k$ can be extended to a bilinear action of $A[[T]]$ on $V[[T]]$ over $k[[T]]$. As before, one could start with a mapping from $A \times V$ into $V[T]$ that is bilinear over $k$, which can be extended to a mapping from $A[[T]] \times V[[T]]$ into $V[[T]]$ that is bilinear over $k[[T]]$.

Suppose now that $k$ is a field with an absolute value function $|\cdot|$, and that $A$, $V$ are vector spaces over $k$ with seminorms $N_A$, $N_V$, respectively, with respect to $|\cdot|$ on $V$. Remember that a bilinear mapping from $A \times V$ into $V$ is said to be bounded with respect to these seminorms if there is a nonnegative real number $C$ such that

\[(6.1.3) \quad N_V(\beta(a, v)) \leq C N_A(a) N_V(v)\]

for every $a \in A$ and $v \in V$, as in Section 1.13. Let $\rho_a$ be as in (6.1.1), so that (6.1.3) can be reformulated as saying that

\[(6.1.4) \quad N_V(\rho_a(v)) \leq C N_A(a) N_V(v)\]

for every $a \in A$ and $v \in V$. This is the same as saying that for each $a \in A$, $\rho_a$ is bounded as a linear mapping from $V$ into itself with respect to $N_V$, with

\[(6.1.5) \quad \|\rho_a\|_{op,V} \leq C N_A(a).\]

This can also be reformulated as saying that $a \mapsto \rho_a$ is bounded as a linear mapping from $A$ into the space $BL(V)$ of bounded linear mappings from $V$ into itself with respect to $N_V$, with the corresponding operator seminorm of this mapping being less than or equal to $C$.

### 6.2 Subactions and homomorphisms

Let $k$ be a commutative ring with a multiplicative identity element, let $A$ and $V$ be modules over $k$, and let $\beta$ be a mapping from $A \times V$ into $V$ that is bilinear over $k$. Also let $W$ be a submodule of $V$, as a module over $k$, and suppose that

\[(6.2.1) \quad \beta(a, w) \in W\]

for every $a \in A$ and $w \in W$. This means that the restriction of $\beta$ to $A \times W$ defines a mapping into $W$ that is bilinear over $k$, and hence a bilinear action of $A$ on $W$. Equivalently, if $\rho_a$ is as in (6.1.1), then (6.2.1) says that

\[(6.2.2) \quad \rho_a(W) \subseteq W\]

for every $a \in A$. Thus the restriction of $\rho_a$ to $W$ defines a module homomorphism from $W$ into itself for every $a \in A$, and the mapping from $a \in A$
6.3. QUOTIENT ACTIONS

to the restriction of $\rho_a$ to $W$ defines a module homomorphism from $A$ into $\text{Hom}_k(W,W)$. If the bilinear action of $A$ on $V$ is expressed as in (6.1.2), then (6.2.1) can be reformulated as saying that

$$(6.2.3) \quad a \cdot w \in W$$

for every $a \in A$ and $w \in W$. If $A$ acts on $V$ on the right, then the corresponding condition is that

$$(6.2.4) \quad w \cdot a \in W$$

for every $a \in A$ and $w \in W$. In this case, $A$ acts on $W$ on the right, as before.

As a basic class of examples, suppose that $A$ is an algebra over $k$ in the strict sense, where multiplication of $a \cdot b \in A$ is expressed as $a b$. We can use multiplication on $A$ to define bilinear actions of $A$ on itself, on the left and on the right. Let $A_0$ be a submodule of $A$, as a module over $k$. The condition that $A_0$ be a left ideal in $A$ says exactly that the action of $A$ on itself on the left maps $A_0$ into itself. Similarly, the condition that $A_0$ be a right ideal in $A$ says that the action of $A$ on itself on the right maps $A_0$ into itself.

Let $A$ and $V$ be modules over $k$ again, and let $\beta^V$ be a mapping from $A \times V$ into $V$ that is bilinear over $k$. Let $Z$ be another module over $k$, let $\beta^Z$ be a mapping from $A \times Z$ into $Z$ that is bilinear over $k$, and let $\phi$ be a homomorphism from $V$ into $Z$, as modules over $k$. If

$$(6.2.5) \quad \phi(\beta^V(a,v)) = \beta^Z(a,\phi(v))$$

for every $a \in A$ and $v \in V$, then we say that $\phi$ intertwines the actions of $A$ on $V$ and $Z$. Equivalently, if $a \in A$, then let $\rho^V_a$ and $\rho^Z_a$ be the module homomorphisms from $V$ and $Z$ into themselves, respectively, associated to $\beta^V$ and $\beta^Z$ as in (6.1.1). It is easy to see that (6.2.5) is the same as saying that

$$(6.2.6) \quad \phi \circ \rho^V_a = \rho^Z_a \circ \phi$$

for every $a \in A$, as mappings from $V$ into $Z$. If these bilinear actions of $A$ on $V$ and $Z$ are expressed as in (6.1.2), then (6.2.5) can be reexpressed as

$$(6.2.7) \quad \phi(a \cdot v) = a \cdot \phi(v)$$

for every $a \in A$ and $v \in V$. If $A$ acts on $V$ and $Z$ on the right, then $\phi$ intertwines these actions when

$$(6.2.8) \quad \phi(v \cdot a) = \phi(v) \cdot a$$

for every $a \in A$ and $v \in V$.

6.3 Quotient actions

Let $k$ be a commutative ring with a multiplicative identity element, let $V$ be a module over $k$, and let $W$ be a submodule of $V$. Remember that the quotient
Suppose that \( \phi \) is a homomorphism from \( V \) into itself, as a module over \( k \), such that
\[
(6.3.1) \quad \phi(W) \subseteq W.
\]
Thus \( q \circ \phi \) is a homomorphism from \( V \) into \( V/W \), as modules over \( k \), whose kernel contains \( W \). Under these conditions, there is a unique mapping \( \psi \) from \( V/W \) into itself such that
\[
(6.3.2) \quad \psi \circ q = q \circ \phi
\]
as mappings from \( V \) into \( W \), by standard arguments. Equivalently, this means
\[
(6.3.3) \quad \psi(q(v)) = q(\phi(v))
\]
for every \( v \in V \). Of course, \( \psi \) is a homomorphism from \( V/W \) into itself, as a module over \( k \).

Let \( A \) be another module over \( k \), and let \( \beta^V \) be a mapping from \( A \times V \) into \( V \) that is bilinear over \( k \). Suppose that the action of \( A \) on \( V \) maps \( W \) into itself, as in (6.2.1), with \( \beta = \beta^V \). This implies that \( q(\beta^V(a,w)) = 0 \) for every \( a \in A \) and \( w \in W \). It follows that for \( a \in A \) and \( v \in V \), \( q(\beta^V(a,v)) \) actually depends only on \( a \) and \( q(v) \). This permits us to define a mapping \( \beta^{V/W} \) from \( A \times (V/W) \) into \( V/W \) that is bilinear over \( k \) and satisfies
\[
(6.3.4) \quad \beta^{V/W}(a,q(v)) = q(\beta^V(a,v))
\]
for every \( a \in A \) and \( v \in V \). Equivalently, if \( a \in A \), then let \( \rho_a^V \) be the module homomorphism from \( V \) into itself associated to \( \beta^V \) as in (6.1.1). There is a unique mapping \( \rho_a^{V/W} \) from \( V/W \) into itself such that
\[
(6.3.5) \quad \rho_a^{V/W} \circ q = q \circ \rho_a^V,
\]
as in (6.3.2). More precisely, \( \rho_a^{V/W} \) is a module homomorphism from \( V/W \) into itself, as before, and one can check that \( a \mapsto \rho_a^{V/W} \) is a module homomorphism from \( A \) into Hom\(_k\)(\( V/W, V/W \)). If the bilinear action of \( A \) on \( V \) is expressed as in (6.1.2), then the induced bilinear action of \( A \) on \( V/W \) can be expressed in the same way, with
\[
(6.3.6) \quad a \cdot q(v) = q(a \cdot v)
\]
for every \( a \in A \) and \( v \in V \). Similarly, if \( A \) acts on \( V \) on the right, then we get a bilinear action of \( A \) on \( V/W \) on the right, with
\[
(6.3.7) \quad q(v) \cdot a = q(v \cdot a)
\]
for every \( a \in A \) and \( v \in V \). Note that \( q \) intertwines the actions of \( A \) on \( V \) and \( V/W \), by construction.

Let \( A \) be an algebra over \( k \) in the strict sense, so that multiplication on \( A \) defines bilinear actions of \( A \) on itself, on the left and on the right, as in the previous section. If \( A_0 \) is a left ideal in \( A \), then the action of \( A \) on itself on the
left maps $A_0$ into itself, and we get a bilinear action of $A$ on the quotient $A/A_0$ on the left, as in the preceding paragraph. Similarly, if $A_0$ is a right ideal in $A$, then the action of $A$ on itself on the right maps $A_0$ into itself, and we get a bilinear action of $A$ on $A/A_0$ on the right. This was mentioned earlier in Section 2.11, in terms of bilinear mappings.

### 6.4 Representations of associative algebras

Let $k$ be a commutative ring with a multiplicative identity element, let $A$ be an associative algebra over $k$, and let $V$ be a module over $k$. Remember that the space $\text{Hom}_k(V, V)$ of module homomorphisms from $V$ into itself is an associative algebra over $k$ with respect to composition of mappings. A representation of $A$ on $V$ is an algebra homomorphism from $A$ into $\text{Hom}_k(V, V)$. If $A$ has a multiplicative identity element $e$, then one may also require that the representation send $e$ to the identity mapping on $V$. In this case, if $a \in A$ has a multiplicative inverse in $A$, then the representation sends $a$ to an invertible mapping on $V$.

A representation of $A$ on $V$ may be denoted $\rho$, where for each $a \in A$, $\rho_a$ denotes the corresponding module homomorphism from $V$ into itself. Thus $\rho_a(v)$ is the image of $v \in V$ under $\rho_a$. Note that $\rho_a(v)$ is linear over $k$ in $a$ and $v$, because the representation is linear over $k$ as a mapping from $A$ into $\text{Hom}_k(V, V)$, and elements of $\text{Hom}_k(V, V)$ are linear over $k$ by definition. The multiplicative property of an algebra homomorphism can be expressed as

\[
\rho_a \circ \rho_b = \rho_{ab}
\]

for every $a, b \in A$, which is the same as saying that

\[
\rho_a(\rho_b(v)) = \rho_{ab}(v)
\]

for every $a, b \in A$ and $v \in V$. If $A$ has a multiplicative identity element $e$, then one may require that $\rho_e$ be the identity mapping on $V$, as before, so that

\[
\rho_e(v) = v
\]

for every $v \in V$.

It is sometimes convenient to express a representation $\rho$ of $A$ on $V$ by

\[
\rho_a(v) = a \cdot v
\]

for every $a \in A$ and $v \in V$. As before, $a \cdot v$ should be linear over $k$ in $a$ and $v$, so that $a \cdot v$ corresponds to a mapping from $A \times V$ into $V$ that is bilinear over $k$. The multiplicative property (6.4.2) can be reexpressed in this notation as

\[
a \cdot (b \cdot v) = (ab) \cdot v
\]

for every $a, b \in A$ and $v \in V$. If $A$ has a multiplicative identity element $e$, then (6.4.3) can be reexpressed as

\[
e \cdot v = v
\]
for every \( v \in V \). We may also call \( V \) a \((\text{left})\) module over \( A \), as an associative algebra over \( k \), with respect to this representation.

Suppose now that we have an action of \( A \) on \( V \) on the right, so that for each \( a \in A \) and \( v \in V \), \( v \cdot a \) is defined as an element of \( V \). Suppose that \( v \cdot a \) is linear over \( k \) in \( a \) and \( v \), so that \( v \cdot a \) corresponds to a mapping from \( V \times A \) into \( V \) that is bilinear over \( k \). If we also have that

\[
(v \cdot a) \cdot b = v \cdot (a b)
\]

for every \( a, b \in A \) and \( v \in V \), then \( V \) is said to be a \((\text{right})\) module over \( A \), as an associative algebra over \( k \). If \( A \) has a multiplicative identity element \( e \), then one may require that

\[
v \cdot e = v
\]

for every \( v \in V \), as usual.

If \( V \) is a right module over \( A \), as in the preceding paragraph, then

\[
(6.4.9) \quad \rho_a(v) = v \cdot a
\]

defines a module homomorphism from \( V \) into itself for each \( a \in A \), because \( v \cdot a \) is linear over \( k \) in \( v \). The mapping from \( a \in A \) to \( \rho_a \in \text{Hom}_k(V, V) \) is linear over \( k \), because \( v \cdot a \) is linear over \( k \) in \( a \). The multiplicativity condition (6.4.7) is the same as saying that

\[
(6.4.10) \quad \rho_b(\rho_a(v)) = \rho_{ab}(v)
\]

for every \( a, b \in A \) and \( v \in V \), which means that

\[
(6.4.11) \quad \rho_a \circ \rho_b = \rho_{ab}
\]

for every \( a, b \in A \). Thus \( a \mapsto \rho_a \) is an opposite algebra homomorphism from \( A \) into \( \text{Hom}_k(V, V) \). If \( A \) has a multiplicative identity element \( e \), then (6.4.8) says that \( \rho_e \) is the identity mapping on \( V \).

Remember that \( A \) is a module over \( k \) in particular. We may also consider \( A \) as both a right and left module over itself, where the actions of \( A \) on itself as a module over \( k \) on the left and the right are given by multiplication on \( A \). The linearity conditions for these actions correspond to the definition of an algebra over \( k \) in the strict sense. Similarly, the conditions (6.4.5) and (6.4.7) correspond in this situation to associativity of multiplication on \( A \). If \( A \) has a multiplicative identity element \( e \), then (6.4.6) and (6.4.8) hold automatically.

Equivalently, the representation of \( A \) on itself as in (6.4.4) corresponds to the multiplication operators discussed in Section 2.2. Similarly, if \( V = A \), then (6.4.9) corresponds to the right multiplication operators discussed in Section 2.7.

### 6.5 Representations of Lie algebras

Let \( k \) be a commutative ring with a multiplicative identity element, let \((A, [\cdot, \cdot]_A)\) be a Lie algebra over \( k \), and let \( V \) be a module over \( k \). The space \( \text{Hom}_k(V, V) \) of
module homomorphisms from $V$ into itself is an associative algebra over $k$ with respect to compositions of mappings, and hence a Lie algebra over $k$ with respect to the corresponding commutator bracket. A Lie algebra homomorphism from $A$ into $\text{Hom}_k(V,V)$, as a Lie algebra over $k$, is also known as a Lie algebra representation of $A$ on $V$.

As before, a Lie algebra representation of $A$ on $V$ may be denoted $\rho$, where $\rho_a$ is the module homomorphism from $V$ into itself corresponding to $a \in A$, and $\rho_a(v)$ is the image of $v \in V$ under $\rho_a$. Thus $\rho_a(v)$ is linear over $k$ in $a$ and $v$, and

\[(6.5.1) \quad \rho_{[a,b]} = \rho_b \circ \rho_a - \rho_a \circ \rho_b \]

for every $a, b \in A$. Equivalently, this means that

\[(6.5.2) \quad \rho_{[a,b]}(v) = \rho_a(\rho_b(v)) - \rho_b(\rho_a(v)) \]

for every $a, b \in A$ and $v \in V$.

Let $\rho$ be a Lie algebra representation of $A$ on $V$, and put

\[(6.5.3) \quad \rho_a(v) = a \cdot v \]

for every $a \in A$ and $v \in V$. This is linear over $k$ in $a$ and $v$, so that it corresponds to a mapping from $A \times V$ into $V$ that is bilinear over $k$. Using this notation, (6.5.2) can be reexpressed as saying that

\[(6.5.4) \quad ([a, b]_A) \cdot v = a \cdot (b \cdot v) - b \cdot (a \cdot v) \]

for every $a, b \in A$ and $v \in V$. We may also call $V$ a module over $A$, as a Lie algebra over $k$, with respect to this representation.

Suppose for the moment that $A$ is an associative algebra over $k$, where multiplication of $a, b \in A$ is expressed as $ab$. Thus $A$ may be considered as a Lie algebra over $k$ with respect to the corresponding commutator bracket $[a, b] = ab - ba$. If $\rho$ is a representation of $A$ on $V$, where $A$ is considered as an associative algebra over $k$, then $\rho$ is a Lie algebra representation of $A$ on $V$ too. Equivalently, if $V$ is a left module over $A$ as an associative algebra over $k$, then $V$ is a module over $A$ as a Lie algebra over $k$ as well. Suppose now that $V$ is a right module over $A$ as an associative algebra over $k$, and let $\rho$ be as in (6.4.9). This means that $a \mapsto \rho_a$ is an opposite algebra homomorphism from $A$ into $\text{Hom}_k(V,V)$, as before. If $a, b \in A$, then

\[(6.5.5) \quad \rho_{[a,b]} = \rho_a \circ \rho_b - \rho_b \circ \rho_a \]

It follows that $-\rho_a$ defines a Lie algebra representation of $A$ on $V$ in this case.

Let $A$ be any Lie algebra over $k$, and remember that the corresponding adjoint representation was defined in Section 2.4. This defines a representation of $A$ on itself, as a module over $k$. 


6.6 Subrepresentations

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an associative algebra over $k$. Also let $\rho$ be a representation of $A$ on a module $V$ over $k$. Suppose that $W$ is a submodule of $V$ such that

$$\rho_a(W) \subseteq W$$

for every $a \in A$. Thus, for each $a \in A$, the restriction of $\rho_a$ to $W$ defines a module homomorphism from $W$ into itself. This defines a representation of $A$ on $W$, which is a subrepresentation of $\rho$ on $V$.

As before, $V$ may be considered as a left module over $A$, with $a \cdot v = \rho_a(v)$ for every $a \in A$ and $v \in V$. Using this notation, (6.6.1) is the same as saying that

$$a \cdot w \in W$$

for every $a \in A$ and $w \in W$. Under these conditions, we may also say that $W$ is a (left) submodule of $V$, as a left module over $A$.

Similarly, suppose that $V$ is a right module over $A$. If $W$ is a submodule of $V$ as a module over $k$, and if

$$w \cdot a \in W$$

for every $a \in A$ and $w \in W$, then we say that $W$ is a (right) submodule of $V$, as a right module over $A$.

As a basic class of examples, let $V$ be a module over $k$, and let $A$ be a subalgebra of $\text{Hom}_k(V, V)$, as an associative algebra over $k$ with respect to composition of mappings. There is an obvious representation of $A$ on $V$, because the elements of $A$ are already module homomorphisms from $V$ into itself. Let $W$ be a submodule of $V$, as a module over $k$, such that

$$a(W) \subseteq W$$

for every $a \in A$. The restrictions of the elements of $A$ to $W$ defines a subrepresentation of the obvious representation of $A$ on $V$ just mentioned. Equivalently, $V$ is a left module over $A$ in an obvious way, and $W$ is a left submodule of $V$ as a left module over $A$.

As another basic class of examples, let $A$ be any associative algebra over $k$. We may consider $A$ as both a left and right module over itself, using multiplication on the left and on the right, as in Section 6.4. A left ideal in $A$ is the same as a left submodule of $A$ as a left module over itself, and similarly a right ideal in $A$ is the same as a right submodule of $A$ as a right module over itself.

Now let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$, and let $\rho$ be a representation of $A$ as a Lie algebra on a module $V$ over $k$. Suppose that $W$ is a submodule of $V$, as a module over $k$, such that (6.6.1) holds for every $a \in A$. Hence the restriction of $\rho_a$ to $W$ defines a homomorphism from $W$ into itself, as a module over $k$, for every $a \in A$. This defines a representation of $A$ as a Lie algebra on $W$, which is a subrepresentation of $\rho$ on $V$. 

As usual, $V$ may be considered as a module over $A$ as a Lie algebra, with $a \cdot v = \rho_a(v)$ for every $a \in A$ and $v \in V$. The condition (6.6.1) on $W$ can be reexpressed in this notation as (6.6.2), as before. In this case, $W$ may be called a submodule of $V$, as a module over $A$, as a Lie algebra over $k$.

Let $V$ be a module over $k$, and let $A$ be a Lie subalgebra of $\text{Hom}_k(V,V)$, as a Lie algebra with respect to the usual commutator bracket. As before, there is an obvious representation of $A$ as a Lie algebra over $k$ on $V$, because the elements of $A$ are already homomorphisms from $V$ into itself, as a module over $k$. If $W$ is a submodule of $V$, as a module over $k$, that satisfies (6.6.4) for every $a \in A$, then the restrictions of the elements of $A$ to $W$ defines a subrepresentation of this representation of $A$ as a Lie algebra over $k$ on $V$. This is the same as saying that $V$ is a module over $A$ as a Lie algebra over $k$ in an obvious way, and that $W$ is a submodule of $V$ as a module over $A$.

If $A$ is any Lie algebra over $k$, then subrepresentations of the adjoint representation of $A$ correspond exactly to ideals in $A$.

### 6.7 Homomorphisms between representations

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an associative algebra over $k$ again. Also let $V$, $W$ be modules over $k$, and let $\rho^V$, $\rho^W$ be representations of $A$ on $V$, $W$, respectively. Suppose that $\phi$ is a homomorphism from $V$ into $W$, as modules over $k$, such that

\begin{equation}
\phi \circ \rho^V_a = \rho^W_a \circ \phi
\end{equation}

for every $a \in A$, which is the same as saying that

\begin{equation}
\phi(\rho^V_a(v)) = \rho^W_a(\phi(v))
\end{equation}

for every $a \in A$ and $v \in V$. In this case, we say that $\phi$ intertwines the representations $\rho^V$ and $\rho^W$, or that $\phi$ is a homomorphism between these representations. If $\phi$ is a one-to-one mapping from $V$ onto $W$, then $\phi^{-1}$ intertwines $\rho^W$ and $\rho^V$ as a mapping from $W$ onto $V$, and we say that $\phi$ defines an isomorphism between $\rho^V$ and $\rho^W$.

Let us consider $V$ and $W$ as left modules over $A$, with

\begin{equation}
a \cdot v = \rho^V_a(v), \quad a \cdot w = \rho^W_a(w)
\end{equation}

for every $a \in A$, $v \in V$, and $w \in W$. Thus (6.7.2) may be reexpressed as

\begin{equation}
\phi(a \cdot v) = a \cdot \phi(v)
\end{equation}

for every $a \in A$ and $v \in V$. We may also say that $\phi$ defines a homomorphism from $V$ into $W$ as left modules over $A$ in this situation. If $\phi$ is a one-to-one mapping from $V$ onto $W$, then it follows that $\phi^{-1}$ is a homomorphism from $W$ into $V$, as left modules over $A$. Under these conditions, $\phi$ is said to be an isomorphism from $V$ onto $W$, as left modules over $A$. 

Suppose now that $V$ and $W$ are right modules over $A$, so that $v \cdot a$ and $w \cdot a$ are defined as elements of $V$ and $W$, respectively, for every $a \in A$, $v \in V$, and $w \in W$. Let $\phi$ be a homomorphism from $V$ into $W$, as modules over $k$. If
\[ \phi(v \cdot a) = \phi(v) \cdot a \]
for every $a \in A$ and $v \in V$, then $\phi$ is said to be a homomorphism from $V$ into $W$, as right modules over $A$. If $\phi$ is also a one-to-one mapping from $V$ onto $W$, then $\phi^{-1}$ is a homomorphism from $W$ into $V$, as right modules over $A$. As before, $\phi$ is said to be an isomorphism from $V$ onto $W$, as right modules over $A$, under these conditions.

Let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$, and let $\rho^V$, $\rho^W$ be representations of $A$ as a Lie algebra on modules $V$, $W$ over $k$, respectively. Also let $\phi$ be a homomorphism from $V$ into $W$, as modules over $k$. If $\phi$ satisfies (6.7.1) for every $a \in A$, then $\phi$ is said to intertwine $\rho^V$ and $\rho^W$, or equivalently be a homomorphism between these representations. We may consider $V$ and $W$ as modules over $A$ as a Lie algebra over $k$, as in (6.7.3). Using this notation, we can reexpress (6.7.2) as (6.7.4), and we say that $\phi$ is a homomorphism from $V$ into $W$ as modules over $A$, as a Lie algebra over $k$. If $\phi$ is a one-to-one mapping from $V$ onto $W$, then $\phi^{-1}$ is a homomorphism from $W$ into $V$, as modules over $A$. In this case, $\phi$ is an isomorphism between these representations of $A$, or equivalently an isomorphism from $V$ onto $W$, as modules over $A$.

### 6.8 Quotient representations

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an associative algebra over $k$. Also let $V$ be a module over $k$, and let $\rho^V$ be a representation of $A$ on $V$. Suppose that
\[ \rho^V_a(W) \subseteq W \]
for every $a \in A$, as in Section 6.6. This implies that for each $a \in A$ there is a unique mapping $\rho^{V/W}_a$ from $V/W$ into itself such that
\[ \rho^{V/W}_a \circ q = q \circ \rho^V_a \]
as mappings from $V$ into $V/W$, as in Section 6.3. One can check that this defines $\rho^{V/W}$ as a representation of $A$ on $V/W$.

Equivalently, suppose that $V$ is a left module over $A$, and that $W$ is a left submodule of $V$. If $a \in A$ and $v \in V$, then $q(a \cdot v)$ defines an element of $V/W$ that is equal to 0 when $v \in W$. This permits us to define an action of $A$ on $V/W$ on the left, with
\[ a \cdot q(v) = q(a \cdot v) \]
for every $a \in A$ and $v \in V$. This defines $V/W$ as a left module over $A$, as before. If $A_0$ is a left ideal in $A$, then note that the quotient $A/A_0$ is a left module over $A$. 
6.9. SUMS OF REPRESENTATIONS

Similarly, suppose that $V$ is a right module over $A$, and that $W$ is a right submodule of $V$. If $a \in A$ and $v \in V$, then $q(v \cdot a)$ is an element of $V/W$ that is equal to 0 when $v \in W$. Using this, we can define an action of $A$ on $V/W$ on the right, with

\begin{equation}
q(v \cdot a) = q(v \cdot a)
\end{equation}

for every $a \in A$ and $v \in V$. One can verify that this defines $V/W$ as a right module over $A$. If $A_0$ is a right ideal in $A$, then the quotient $A/A_0$ is a right module over $A$.

Suppose now that $A$ is a Lie algebra over $k$, and that $\rho^V$ is a representation of $A$ on $V$. If $W$ satisfies (6.8.1) for every $a \in A$, then one can define $\rho^{V/W}$ on $V/W$ as in (6.8.2). One can check that $\rho^{V/W}$ is a representation of $A$ on $V/W$. Equivalently, if $V$ is a module over $A$, and if $W$ is a submodule of $V$ as a module over $A$, then we can define the action of $A$ on $V/W$ as in (6.8.3). This makes $V/W$ a module over $A$, as a Lie algebra over $k$, as before.

In each of these situations, the quotient mapping $q$ intertwines the actions of $A$ on $V$ and $V/W$, by construction.

6.9 Sums of representations

Let $k$ be a commutative ring with a multiplicative identity element, and let $I$ be a nonempty set. Also let $V_j$ be a module over $k$ for every $j \in I$. Thus the direct product $\prod_{j \in I} V_j$ of the $V_j$’s defines a module over $k$, and the direct sum $\bigoplus_{j \in I} V_j$ is a submodule of $\prod_{j \in I} V_j$, as in Section 4.1. If $v \in \prod_{j \in I} V_j$ and $j \in I$, then $v_j$ denotes the $j$th coordinate of $v$ in $V_j$, as before.

Let $A$ be a module over $k$, and suppose that for each $j \in I$, $\beta^j$ is a mapping from $A \times V_j$ into $V_j$ that is bilinear over $k$. If $a \in A$ and $v \in \prod_{j \in I} V_j$, then $\beta(a, v)$ can be defined as an element of $\prod_{j \in I} V_j$ by putting

\begin{equation}
(\beta(a, v))_j = \beta^j(a, v_j)
\end{equation}

for every $j \in I$. It is easy to see that this defines $\beta$ as a mapping from $A \times \prod_{j \in I} V_j$ into $\prod_{j \in I} V_j$ that is bilinear over $k$. If $v \in \bigoplus_{j \in I} V_j$, then

\begin{equation}
\beta(a, v) \in \bigoplus_{j \in I} V_j
\end{equation}

for every $a \in A$.

Equivalently, if the bilinear action of $A$ on $V_j$ is given by $\rho^j$ for each $j \in I$, as in Section 6.1, then the corresponding bilinear action $\rho$ of $A$ on $\prod_{j \in I} V_j$ is defined by

\begin{equation}
(\rho_a(v))_j = \rho^j_a(v_j).
\end{equation}

Similarly, if the bilinear action of $A$ on $V_j$ is expressed as $a \cdot v_j$ for every $a \in A$, $v_j \in V_j$, and $j \in I$, then the corresponding bilinear action of $A$ on $\prod_{j \in I} V_j$ can be expressed by $a \cdot v$ for every $a \in A$ and $v \in \prod_{j \in I} V_j$, where

\begin{equation}
(a \cdot v)_j = a \cdot v_j
\end{equation}
for each \( j \in I \). If \( A \) acts on \( V_j \) on the right for each \( j \in I \), so that the bilinear action is expressed as \( v_j \cdot a \) for every \( a \in A \), \( v_j \in V_j \), and \( j \in I \), then the corresponding bilinear action of \( A \) on \( \prod_{j \in I} V_j \) on the right can be expressed as \( v \cdot a \) for every \( a \in A \) and \( v \in \prod_{j \in I} V_j \), where
\[
(6.9.5) \quad (v \cdot a)_j = v_j \cdot a
\]
for each \( j \in I \).

Let \( A \) be an associative algebra over \( k \). If \( \rho^V \) is a representation of \( A \) on \( V_j \) for each \( j \in I \), then one can check that (6.9.3) defines \( \rho \) as a representation of \( A \) on \( \prod_{j \in I} V_j \). Equivalently, if \( V_j \) is a left module over \( A \) for every \( j \in I \), then \( \prod_{j \in I} V_j \) is a left module over \( A \) with respect to (6.9.4). Note that \( \bigoplus_{j \in I} V_j \) is a left submodule of \( \prod_{j \in I} V_j \), as a left module over \( A \), as in (6.9.2). Similarly, if \( V_j \) is a right module over \( A \) for every \( j \in I \), then \( \prod_{j \in I} V_j \) is a right module over \( A \) with respect to (6.9.5), and \( \bigoplus_{j \in I} V_j \) is a right submodule of \( \prod_{j \in I} V_j \), as a right module over \( A \).

Now let \( A \) be a Lie algebra over \( k \). If \( \rho^V \) is a representation of \( A \) on \( V_j \) for every \( j \in I \), then one can verify that (6.9.3) defines \( \rho \) as a representation of \( A \) on \( \prod_{j \in I} V_j \). Equivalently, if \( V_j \) is a module over \( A \) for every \( j \in I \), then \( \prod_{j \in I} V_j \) is a module over \( A \) with respect to (6.9.4). As before, \( \bigoplus_{j \in I} V_j \) is a submodule of \( \prod_{j \in I} V_j \), as a module over \( A \).

Let \( V \) be a module over \( k \), and let \( V_1, V_2 \) be submodules of \( V \). Observe that
\[
(6.9.6) \quad V_1 + V_2 = \{ v_1 + v_2 : v_1 \in V_1, \ v_2 \in V_2 \}
\]
is a submodule of \( V \) too. The direct sum \( V_1 \oplus V_2 \) of \( V_1 \) and \( V_2 \) can be defined as a module over \( k \) as in Section 4.1, with \( I = \{1, 2\} \). More precisely, \( V_1 \oplus V_2 \) can be defined as a set as the Cartesian product \( V_1 \times V_2 \) of \( V_1 \) and \( V_2 \), consisting of all ordered pairs \((v_1, v_2)\) with \( v_1 \in V_1 \) and \( v_2 \in V_2 \). Addition and scalar multiplication can be defined on \( V_1 \oplus V_2 \) coordinatewise, as usual. Observe that
\[
(6.9.7) \quad (v_1, v_2) \mapsto v_1 + v_2
\]
defines a homomorphism from \( V_1 \oplus V_2 \) onto \( V_1 + V_2 \), as modules over \( k \). If
\[
(6.9.8) \quad V_1 \cap V_2 = \{0\},
\]
then (6.9.7) is injective as a mapping from \( V_1 \oplus V_2 \) into \( V \).

Let \( A \) be a module over \( k \) again, and suppose that we have a bilinear action of \( A \) on \( V \). If this action maps \( V_1 \) and \( V_2 \) into themselves, then it maps \( V_1 + V_2 \) into itself as well. In this case, we also get a corresponding bilinear action of \( A \) on \( V_1 \oplus V_2 \), as before. Of course, (6.9.7) intertwines the bilinear actions of \( A \) on \( V_1 \oplus V_2 \) and \( V_1 + V_2 \).

### 6.10 Compatible bilinear mappings

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( A \) and \( V \) be modules over \( k \). Also let \( \rho \) be a bilinear action of \( A \) on \( V \), so that
6.10. COMPATIBLE BILINEAR MAPPINGS

$\rho_a$ is a module homomorphism from $V$ into itself for every $a \in A$, and $a \mapsto \rho_a$ is a module homomorphism from $A$ into $\text{Hom}_k(V,V)$, as in Section 6.1. Let $W$ be another module over $k$, and let $\mu$ be a mapping from $V \times V$ into $W$ that is bilinear over $k$. Suppose that for every $a \in A$ there is an $\tilde{a} \in A$ such that

\[(6.10.1) \quad \mu(\rho_a(u), v) = \mu(u, \rho_{\tilde{a}}(v))\]

for every $u, v \in V$. If $k = \mathbb{C}$, then one may consider a sesquilinear form $\mu$ on $V$.

Let $V_0$ be a submodule of $V$, and put

\[(6.10.2) \quad (V_0)^\perp = \{ u \in V : \text{ for every } v \in V_0, \mu(u, v) = 0 \}.\]

Suppose that $\rho$ satisfies the compatibility condition with $\mu$ in the preceding paragraph, and that

\[(6.10.3) \quad \rho_a(V_0) \subseteq V_0\]

for every $a \in A$. Let $a \in A$, $u \in (V_0)^\perp$, and $v \in V$ be given, and let $\tilde{a} \in A$ be as in (6.10.1). Observe that

\[(6.10.4) \quad \mu(\rho_a(u), v) = \mu(u, \rho_{\tilde{a}}(v)) = 0,\]

because $\rho_{\tilde{a}}(v) \in V_0$, by (6.10.3). This means that $\rho_a(u) \in (V_0)^\perp$, so that

\[(6.10.5) \quad \rho_a((V_0)^\perp) \subseteq (V_0)^\perp.\]

Now let $A$ be an associative algebra over $k$, and suppose that $\rho$ is a representation of $A$ on $V$. Also let $a \mapsto a^*$ be an opposite algebra automorphism on $A$. One way that (6.10.1) can hold is with $\tilde{a} = a^*$, so that

\[(6.10.6) \quad \mu(\rho_a(u), v) = \mu(u, \rho_{a^*}(v))\]

for every $a \in A$ and $u, v \in V$. If $k = \mathbb{C}$ and $\mu$ is a sesquilinear form on $V$, then one may consider a conjugate-linear opposite algebra automorphism on $A$.

If $\rho$ satisfies (6.10.6), then

\[(6.10.7) \quad \mu(\rho_a(u), \rho_a(v)) = \mu(u, \rho_{a^*}(\rho_a(v))) = \mu(u, \rho_{a^*}a(v))\]

for every $a \in A$ and $u, v \in V$. Suppose that $A$ has a multiplicative identity element $e$, and that $\rho_e$ is the identity mapping on $V$. If $a \in A$ satisfies $a^* a = e$, then (6.10.7) implies that $\rho_a$ preserves $\mu$.

In some situations there may be an opposite algebra automorphism $T \mapsto T^*$ on the algebra $\text{Hom}_k(V,V)$ of module homomorphisms from $V$ into itself such that

\[(6.10.8) \quad \mu(T(u), v) = \mu(u, T^*(v))\]

for every $u, v \in V$ and $T \in \text{Hom}_k(V,V)$. In this case, one can ask that $\rho$ be compatible with these opposite algebra automorphisms on $A$ and $\text{Hom}_k(V,V)$, in the sense that

\[(6.10.9) \quad (\rho_a)^* = \rho_{a^*}.\]
for every \( a \in A \). Note that (6.10.6) follows from (6.10.8) and (6.10.9). More precisely, one might have an opposite algebra automorphism on a subalgebra of \( \text{Hom}_k(V,V) \) that satisfies (6.10.8), and a representation \( \rho \) of \( A \) on \( V \) with values in this subalgebra of \( \text{Hom}_k(V,V) \). If \( k = \mathbb{C} \), then one may consider conjugate-linear opposite algebra automorphisms and a sesquilinear form on \( V \) again.

Suppose now that \((A, \cdot, A)\) is a Lie algebra over \( k \), and that \( \rho \) is a representation of \( A \) as a Lie algebra on \( V \). A natural compatibility condition for \( \rho \) is that \( \rho \) be antisymmetric with respect to \( \rho \) for every \( a \in A \), so that

\[
\mu(\rho_a(u), v) = -\mu(u, \rho_a(v))
\]

(6.10.10)

for every \( u, v \in V \). This means that (6.10.1) holds with \( \tilde{a} = -a \).

Let us take \( V = A \), as a module over \( k \), and \( \rho \) to be the adjoint representation on \( A \). Thus, for each \( x \in A \), \( \rho_x = \text{ad} x = \text{ad}_x \) is the module homomorphism from \( A \) into itself defined by

\[
\text{ad}_x(z) = [x, z]_A,
\]

(6.10.11)

as in Section 2.4. In this situation, (6.10.10) is the same as saying that

\[
\mu(\text{ad}_w(x), y) = -\mu(x, \text{ad}_w(y))
\]

(6.10.12)

for every \( w, x, y \in A \). Equivalently, this means that

\[
\mu([w, x]_A, y) = -\mu(x, [w, y]_A)
\]

(6.10.13)

for every \( w, x, y \in A \), which can also be expressed as

\[
\mu([x, w]_A, y) = \mu(x, [w, y]_A).
\]

(6.10.14)

This property is sometimes described by saying that \( \mu \) is associative on \( A \times A \), as on p21 of [13].

### 6.11 Representations and formal power series

Let \( k \) be a commutative ring with a multiplicative identity element, let \( V \) be a module over \( k \), and let \( T \) be an indeterminate. Remember that there are natural isomorphisms between \( \text{Hom}_k(V, V[[T]]) \),

\[
\text{Hom}_k(V[[T]], V[[T]])
\]

(6.11.1)

and

\[
(\text{Hom}_k(V, V))[T],
\]

(6.11.2)

as modules over \( k[[T]] \), as in Sections 4.8 and 4.9. Of course, \( \text{Hom}_k(V, V) \) and (6.11.1) are associative algebras over \( k \) and \( k[[T]] \), respectively, with composition of mappings as multiplication. Similarly, (6.11.2) is an associative algebra over
6.12. OPPOSITE ALGEBRAS

Let $k[[T]]$, as in Section 6.4. In fact, the natural isomorphism between (6.11.2) and (6.11.1) preserves multiplication, as in Section 4.10.

As before, there is also a natural isomorphism between $\text{Hom}_k(V,V[[T]])$ and

$$\text{Hom}_{k[[T]]}(V[[T]],V[[T]])$$

and a natural injective homomorphism from

$$\text{Hom}_k(V,V)[[T]]$$

into (6.11.3), as modules over $k[[T]]$. Note that (6.11.3) is an associative algebra over $k[[T]]$ with respect to composition of mappings, and that (6.11.4) is an associative algebra over $k[[T]]$ too, as in Section 6.4. The natural injective homomorphism from (6.11.4) into (6.11.3) preserves multiplication, as in Section 4.10.

Let $A$ be an associative algebra over $k$, so that $A[[T]]$ and $A[T]$ are associative algebras over $k[[T]]$ and $k[T]$, respectively, as in Section 4.6. Remember that a representation of $A$ on $V$ corresponds to an algebra homomorphism from $A$ into $\text{Hom}_k(V,V)$, as in Section 6.4. This can be extended to an algebra homomorphism from $A[[T]]$ into (6.11.2), as algebras over $k[[T]]$, as in Section 4.12. This corresponds to an algebra homomorphism from $A[[T]]$ into (6.11.1), which is to say a representation of $A[[T]]$ on $V[[T]]$. Similarly, we get an algebra homomorphism from $A[T]$ into (6.11.4), as algebras over $k[T]$, as in Section 4.12. This leads to an algebra homomorphism from $A[T]$ into (6.11.3), which gives a representation of $A[T]$ on $V[[T]]$. If $V$ is a right module over $A$, then we can get $V[[T]]$ as a right module over $A[[T]]$, and $V[T]$ as a right module over $A[T]$, in the same way.

Remember that an associative algebra is automatically a Lie algebra with respect to the corresponding commutator bracket. The natural algebra isomorphism between (6.11.2) and (6.11.1) mentioned earlier automatically preserves commutator brackets. Similarly, the natural injective algebra homomorphism from (6.11.4) into (6.11.3) automatically preserves commutator brackets. Let $A$ be a Lie algebra over $k$, so that $A[[T]]$ and $A[T]$ are Lie algebras over $k[[T]]$ and $k[T]$, as in Section 4.6. A representation of $A$ on $V$ is the same as a homomorphism from $A$ into $\text{Hom}_k(V,V)$, as a Lie algebra over $k$ with respect to the commutator bracket. This can be extended to a Lie algebra homomorphism from $A[[T]]$ into (6.11.2), as in Section 4.12. This corresponds to a Lie algebra homomorphism from $A[[T]]$ into (6.11.1), which is a representation of $A[[T]]$ as a Lie algebra over $k[[T]]$ on $V[[T]]$. We also get a Lie algebra homomorphism from $A[T]$ into (6.11.4), as in Section 4.12. This leads to a Lie algebra homomorphism from $A[T]$ into (6.11.3), and hence a representation of $A[T]$ as a Lie algebra over $k[T]$ on $V[T]$.

6.12 Opposite algebras

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an algebra over $k$ in the strict sense, where multiplication of $a, b \in A$ is
expressed as $a b$. The corresponding opposite algebra $A^{op}$ is defined as an algebra over $k$ in the strict sense as follows. As a module over $k$, $A^{op}$ is the same as $A$. The product of $a, b \in A^{op}$ is defined to be the product $ba$ of $b$ and $a$ in $A$. Thus multiplication in $A^{op}$ is the same as multiplication in $A$ exactly when $A$ is commutative. If $A$ has a multiplicative identity element $e$, then $e$ is also the multiplicative identity element in $A^{op}$. If $A$ is associative, then it is easy to see that $A^{op}$ is associative as well. By construction, the identity mapping on $A$ is an opposite algebra isomorphism between $A$ and $A^{op}$.

Let $B$ be another algebra over $k$ in the strict sense. An algebra homomorphism from $A$ into $B$ may also be considered as an algebra homomorphism from $A^{op}$ into $B^{op}$. An opposite algebra homomorphism from $A$ into $B$ corresponds to an algebra homomorphism from $A^{op}$ into $B$, or from $A$ into $B^{op}$.

Suppose that $A$ is an associative algebra over $k$, and let $V$ be a module over $k$. A bilinear action of $A$ on $V$ makes $V$ into a left module over $A$ exactly when it makes $V$ into a right module over $A^{op}$. Similarly, a bilinear action of $A$ on $V$ makes $V$ into a right module over $A$ exactly when it makes $V$ into a left module over $A^{op}$. Equivalently, this corresponds to a representation of $A^{op}$ on $V$.

Let $n$ be a positive integer, and let $M_n(A)$ and $M_n(A^{op})$ be the corresponding spaces of $n \times n$ matrices with entries in $A$ and $A^{op}$, respectively, as in Section 2.8. Thus $M_n(A)$ and $M_n(A^{op})$ are the same as modules over $k$, using entrywise addition and scalar multiplication. Remember that $M_n(A)$ and $M_n(A^{op})$ are algebras over $k$ with respect to matrix multiplication. Let $a, b \in M_n(A)$ with entries in $A$, which can also be considered as $n \times n$ matrices with entries in $A^{op}$. The product $c$ of $a$ and $b$ in $M_n(A)$ is given by

$$
(6.12.1) 
\quad c_{j,r} = \sum_{l=1}^{n} a_{j,l} b_{l,r} 
$$

for every $j, r = 1, \ldots, n$, as usual. Let $\bar{c}$ be the product of $a$ and $b$ in $M_n(A^{op})$. This means that

$$
(6.12.2) 
\quad \bar{c}_{j,r} = \sum_{l=1}^{n} b_{l,r} a_{j,l} 
$$

for every $j, r = 1, \ldots, n$, where the terms in the sum on the right use multiplication in $A$.

Let $a^t, b^t$, and $c^t$ be the transposes of $a, b$, and $c$, respectively, as in Section 2.8. Thus

$$
(6.12.3) 
\quad c^t_{j,r} = c_{r,j} = \sum_{l=1}^{n} a_{r,l} b_{l,j} 
$$

for every $j, r = 1, \ldots, n$. This is the same as the product of $b^t$ and $a^t$ in $M_n(A^{op})$. This means that

$$
(6.12.4) 
\quad a \mapsto a^t 
$$

is an opposite algebra homomorphism from $M_n(A)$ into $M_n(A^{op})$. More precisely, (6.12.4) is an opposite algebra isomorphism from $M_n(A)$ onto $M_n(A^{op})$, because (6.12.4) is a one-to-one mapping from $M_n(A)$ onto $M_n(A^{op})$. 


6.13 Matrices and associative algebras

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( A \) be an associative algebra over \( k \). Also let \( n \) be a positive integer, and let \( A^n \) be the space of \( n \)-tuples of elements of \( A \), as usual. Of course, \( A^n \) is a module over \( k \) with respect to coordinatewise addition and scalar multiplication. Similarly, \( A^n \) may be considered as both a left and right module over \( A \), with respect to coordinatewise multiplication. More precisely, if \( a \in A \) and \( x = (x_1, \ldots, x_n) \) is an element of \( A^n \), then \( a \cdot x \) and \( x \cdot a \) are defined as elements of \( A^n \) by

\[
(6.13.1) \quad a \cdot x = (a x_1, \ldots, a x_n)
\]

and

\[
(6.13.2) \quad x \cdot a = (x_1 a, \ldots, x_n a),
\]

respectively. If \( A \) is commutative, then (6.13.1) and (6.13.2) are the same. Note that

\[
(a \cdot x) \cdot b = a \cdot (x \cdot b) \quad (6.13.3)
\]

for every \( a, b \in A \) and \( x \in A^n \).

Let \( \alpha = (\alpha_{j,l}) \) be an \( n \times n \) matrix with entries in \( A \), which is to say an element of \( M_n(A) \). If \( x \in A^n \), then let \( T^L_\alpha(x) \) be the element of \( A^n \) whose \( j \)th coordinate is given by

\[
(6.13.4) \quad (T^L_\alpha(x))_j = \sum_{l=1}^{n} \alpha_{j,l} x_l
\]

for every \( j = 1, \ldots, n \). Similarly, let \( T^R_\alpha(x) \) be the element of \( A^n \) whose \( j \)th coordinate is given by

\[
(6.13.5) \quad (T^R_\alpha(x))_j = \sum_{l=1}^{n} x_l \alpha_{j,l}
\]

for every \( j = 1, \ldots, n \). If \( A \) is commutative, then (6.13.4) and (6.13.5) are the same. It is easy to see that \( T^L_\alpha \) and \( T^R_\alpha \) are homomorphisms from \( A^n \) into itself, as a module over \( k \). Observe that

\[
(6.13.6) \quad (T^L_\alpha(x)) \cdot a = T^L_\alpha(x \cdot a)
\]

and

\[
(6.13.7) \quad a \cdot (T^R_\alpha(x)) = T^R_\alpha(a \cdot x)
\]

for every \( a \in A \) and \( x \in A^n \). Thus \( T^L_\alpha \) is a homomorphism from \( A^n \) into itself as a right module over \( A \), and \( T^R_\alpha \) is a homomorphism from \( A^n \) into itself as a left module over \( A \).

If \( t \in k \), then \( t \alpha = (t \alpha_{j,l}) \in M_n(A) \), as in Section 2.8. Clearly

\[
(6.13.8) \quad T^L_{t\alpha}(x) = t T^L_\alpha(x)
\]

and

\[
(6.13.9) \quad T^R_{t\alpha}(x) = t T^R_\alpha(x)
\]
for every $x \in A^n$. Let $\beta = (\beta_{j,l})$ be another $n \times n$ matrix with entries in $A$, so that $\alpha + \beta \in M_n(A)$ too. Of course,

$$T_{\alpha+\beta}^L(x) = T_\alpha^L(x) + T_\beta^L(x)$$

(6.13.10)

and

$$T_{\alpha+\beta}^R(x) = T_\alpha^R(x) + T_\beta^R(x)$$

(6.13.11)

for every $x \in A^n$. One can check that

$$T_L(T_L(x)) = T_L(x)$$

(6.13.12)

for every $x \in A^n$, where $\alpha, \beta \in M_n(A)$ is defined using matrix multiplication, as in Section 2.8. Let $\gamma$ be the product of $\alpha$ and $\beta$ as elements of $M_n(A^{op})$, where $A^{op}$ is the opposite algebra associated to $A$, as in the previous section. One can verify that

$$T_R(T_R(x)) = T_R(x)$$

(6.13.13)

for every $x \in A^n$.

Suppose that $A$ has a multiplicative identity element $e$. Remember that the corresponding identity matrix in $M_n(A)$ has diagonal entries equal to $e$ and all other entries equal to 0, as in Section 2.8. If $\alpha$ is the identity matrix, then $T_\alpha^L$ and $T_\alpha^R$ are equal to the identity mapping on $A^n$. Let $u^1, \ldots, u^n$ be the elements of $A^n$ with $u^j_l = e$ when $j = l$ and $u^j_l = 0$ when $j \neq l$. Thus

$$x = \sum_{l=1}^{n} x_l \cdot u^l = \sum_{l=1}^{n} u^l \cdot x_l$$

(6.13.14)

for every $x \in A^n$.

If $T$ is any homomorphism from $A^n$ into itself, as a right module over $A$, then

$$T(x) = T\left(\sum_{l=1}^{n} u^l \cdot x_l\right) = \sum_{l=1}^{n} T(u^l) \cdot x_l$$

(6.13.15)

for every $x \in A^n$. This means that $T$ can be represented in a unique way as $T_{\alpha}^L$, with $\alpha \in M_n(A)$. More precisely, $\alpha_{j,l}$ is the $j$th coordinate of $T(u^l)$ for each $j, l = 1, \ldots, n$. Similarly, if $T$ is a homomorphism from $A^n$ into itself, as a left module over $A$, then

$$T(x) = T\left(\sum_{l=1}^{n} x_l \cdot u^l\right) = \sum_{l=1}^{n} x_l \cdot T(u^l)$$

(6.13.16)

for every $x \in A^n$. This implies that $T$ can be represented in a unique way as $T_{\alpha}^R$, where $\alpha_{j,l}$ is the $j$th coordinate of $T(u^l)$ for every $j, l = 1, \ldots, n$. 
6.14 Irreducibility

Let $k$ be a field, let $A$ and $V$ be vector spaces over $k$, and let $\rho$ be a bilinear action of $A$ on $V$. Thus $\rho_a$ is a linear mapping from $V$ into itself for every $a \in A$, and $a \mapsto \rho_a$ is a linear mapping from $A$ into the space $\mathcal{L}(V)$ of linear mappings from $V$ into itself, as in Section 6.1. Suppose that there is no linear subspace $W$ of $V$ such that $W \neq \{0\}$, $V$ and

\[(6.14.1) \quad \rho_a(W) \subseteq W\]

for every $a \in A$. In this case, one may say that $\rho$ is irreducible on $V$, or equivalently that $V$ is simple with respect to the action of $\rho$. The condition that $V \neq \{0\}$ is typically included in the definition of irreducibility or simplicity as well.

Let $V_1, V_2$ be vector spaces over $k$, and let $\rho^1, \rho^2$ be bilinear actions of $A$ on $V_1, V_2$, respectively. Suppose that $\phi$ is a linear mapping from $V_1$ into $V_2$ that intertwines $\rho^1$ and $\rho^2$, as in Section 6.2. If $v_1 \in V_1$ is in the kernel of $\phi$ and $a \in A$, then

\[(6.14.2) \quad \phi(\rho^1_a(v_1)) = \rho^2_a(\phi(v_1)) = 0,\]

so that $\rho^2_a(v_1)$ is in the kernel of $\phi$ too. If $\rho^1$ is irreducible on $V_1$, then it follows that the kernel of $\phi$ is either trivial or equal to $V_1$, so that $\phi$ is either injective or equal to 0 on $V_1$. This is part of Schur’s lemma.

Similarly,

\[(6.14.3) \quad \rho^1_a(\phi(V_1)) = \phi(\rho^1_a(V_1)) \subseteq \phi(V_1)\]

for every $a \in A$. If $\rho^2$ is irreducible on $V_2$, then it follows that $\phi(V_1) = \{0\}$ or $V_2$, so that either $\phi = 0$ on $V_1$ or $\phi$ is surjective. This is another part of Schur’s lemma. If $\rho^1$ and $\rho^2$ are both irreducible, then either $\phi = 0$ or $\phi$ is a bijection.

Let $V$ be a vector space over $k$ again, and let $\rho$ be a bilinear action of $A$ on $V$. Remember that $\mathcal{L}(V)$ is an associative algebra over $k$ with respect to composition of mappings. Consider the space $\mathcal{L}(V)$ of $\phi \in \mathcal{L}(V)$ that intertwine $\rho$. It is easy to see that $\mathcal{L}(V)$ is a subalgebra of $\mathcal{L}(V)$, and that $\mathcal{L}(V)$ contains the identity mapping on $V$. If $\rho$ is irreducible, then every nonzero element of $\mathcal{L}(V)$ is invertible, as in the preceding paragraph.

Suppose that $k$ is algebraically closed, and that $V$ has positive finite dimension as a vector space over $k$. If $\phi$ is any linear mapping from $V$ into itself, then it is well known that there is a $\lambda \in k$ such that $\phi$ has a nonzero eigenvector in $V$ with eigenvalue $\lambda$. Let $E_\lambda$ be the corresponding eigenspace of eigenvectors of $\phi$ in $V$ with eigenvalue $\lambda$. Let $\rho$ be a bilinear action of $A$ on $V$ again, and suppose that $\phi$ intertwines $\rho$. If $a \in A$, then $\rho_a$ maps $E_\lambda$ into itself, by a standard argument. If $\rho$ is irreducible on $V$, then it follows that $E_\lambda = V$. This is another part of Schur’s lemma.
Chapter 7

Representations and multilinear mappings

7.1 Some remarks about subalgebras

Let $k$ be a commutative ring with a multiplicative identity element, and let $A_1$, $A_2$ be algebras over $k$ in the strict sense. The direct sum $A_1 \oplus A_2$ of $A_1$ and $A_2$ can be defined as an algebra over $k$ in the strict sense as in Section 4.1, with $I = \{1, 2\}$. More precisely, $A_1 \oplus A_2$ can be defined as a set as the Cartesian product of $A_1$ and $A_2$, consisting of all ordered pairs $(a_1, a_2)$ with $a_1 \in A_1$ and $a_2 \in A_2$. Addition, scalar multiplication, and multiplication on $A_1 \oplus A_2$ are defined coordinatewise, as usual. In particular, if multiplication of $a_1 b_1$ and $a_2 b_2$, respectively, then

$$(a_1, 0) (0, a_2) = 0$$

in $A_1 \oplus A_2$ for every $a_1 \in A_1$ and $a_2 \in A_2$.

If $A$ is a module over $k$ and $A_1, A_2 \subseteq A$ are submodules of $A$, then

$$(7.1.2) \quad A_1 + A_2 = \{a_1 + a_2 : a_1 \in A_1, a_2 \in A_2\}$$

is a submodule of $A$, as before. Suppose that $A$ is an algebra over $k$ in the strict sense. If $A_1$, $A_2$ are left ideals in $A$, then $(7.1.2)$ is a left ideal in $A$. Similarly, if $A_1$, $A_2$ are right ideals in $A$, then $(7.1.2)$ is a right ideal in $A$. If $A_1$, $A_2$ are two-sided ideals in $A$, then it follows that $(7.1.2)$ is a two-sided ideal in $A$.

Let $A_1$, $A_2$ be subalgebras of $A$. If

$$(7.1.3) \quad a_1 a_2 = a_2 a_1 = 0$$

for every $a_1 \in A_1$ and $a_2 \in A_2$, then it is easy to see that $(7.1.2)$ is a subalgebra of $A$. In this situation,

$$(7.1.4) \quad (a_1, a_2) \mapsto a_1 + a_2$$

142
defines a homomorphism from \( A_1 \oplus A_2 \) into \( A \), as algebras over \( k \) in the strict sense. If \((A, [\cdot,\cdot])_A\) is a Lie algebra over \( k \), then (7.1.3) is the same as saying that
\[
[a_1, a_2]_A = 0
\]
for every \( a_1 \in A_1 \) and \( a_2 \in A_2 \).

Let \( A \) and \( B \) be algebras over \( k \) in the strict sense, and let \( \phi, \psi \) be algebra homomorphisms from \( A \) into \( B \). Suppose that
\[
\phi(x) \psi(y) = \psi(y) \phi(x) = 0
\]
for every \( x, y \in A \). This implies that
\[
[\phi(x), \psi(y)]_B = 0
\]
for every \( x, y \in A \), so that \( \phi + \psi \) defines an algebra homomorphism from \( A \) into \( B \) as well. If \((B, [\cdot,\cdot])_B\) is a Lie algebra over \( k \), then (7.1.6) is the same as saying that
\[
[\phi(x), \psi(y)]_B = 0
\]
for every \( x, y \in A \).

Let \( A \) be a Lie algebra over \( k \), and let \( \rho^1, \rho^2 \) be Lie algebra representations of \( A \) on a module \( V \) over \( k \). Let us say that that \( \rho^1 \) and \( \rho^2 \) are commuting representations of \( A \) on \( V \) if
\[
\rho^1_a \circ \rho^2_b = \rho^2_b \circ \rho^1_a
\]
for every \( a, b \in A \). Under these conditions,
\[
(\rho^1 + \rho^2)_a = \rho^1_a + \rho^2_a
\]
defines a Lie algebra representation of \( A \) on \( V \) too, as in the preceding paragraph.

### 7.2 Representations on linear mappings

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( V, W \) be modules over \( k \). Remember that the space \( \text{Hom}_k(V, W) \) of module homomorphisms from \( V \) into \( W \) is a module over \( k \) too, with respect to pointwise addition and scalar multiplication. Let \( T \) be a module homomorphism from \( V \) into itself, and let \( \phi \) be a module homomorphism from \( V \) into \( W \). Thus
\[
\widetilde{T}(\phi) = \phi \circ T
\]
defines another module homomorphism from \( V \) into \( W \). This defines \( \widetilde{T} \) as a homomorphism from \( \text{Hom}_k(V, W) \) into itself, as a module over \( k \).
Let $R$ be another module homomorphism from $V$ into itself, so that the composition $T \circ R$ is a module homomorphism from $V$ into itself as well. In particular, $\widetilde{R}$ and $(T \circ R)$ can be defined as homomorphisms from $\text{Hom}_k(V, W)$, as a module over $k$, as in the preceding paragraph. If $\phi$ is a module homomorphism from $V$ into $W$, then

$$\text{(7.2.2) } (T \circ R)(\phi) = \phi \circ (T \circ R) = (\phi \circ T) \circ R = \widetilde{R}(\widetilde{T}(\phi)).$$

This means that

$$\text{(7.2.3) } (T \circ R) = \widetilde{R} \circ \widetilde{T}$$

as mappings from $\text{Hom}_k(V, W)$. Note that $\widetilde{R}$ is linear over $k$, as a mapping from $\text{Hom}_k(V, V)$ into the space of mappings from $\text{Hom}_k(V, W)$ into itself. More precisely, this defines an opposite algebra homomorphism from $\text{Hom}_k(V, V)$ into the algebra of module homomorphisms from $\text{Hom}_k(V, W)$ into itself. If $T$ is the identity mapping on $V$, then $\widetilde{T}$ is the identity mapping on $\text{Hom}_k(V, W)$.

Let $A$ be an associative algebra over $k$, and let $\rho$ be a representation of $A$ on $V$. If $a \in A$, then $\rho_a$ is a module homomorphism from $V$ into itself, so that $\widetilde{\rho_a}$ can be defined as before, as a module homomorphism from $\text{Hom}_k(V, W)$ into itself. By hypothesis, $a \mapsto \rho_a$ is an algebra homomorphism from $A$ into $\text{Hom}_k(V, V)$, which implies that $\widetilde{\rho_a}$ is an opposite algebra homomorphism from $A$ into the algebra of module homomorphisms from $\text{Hom}_k(V, W)$ into itself. Similarly, if $a \mapsto \widetilde{\rho_a}$ is an algebra homomorphism from $A$ into $\text{Hom}_k(V, V)$, then $a \mapsto \widetilde{\rho_a}$ is an algebra homomorphism from $A$ into the algebra of module homomorphisms from $\text{Hom}_k(V, W)$ into itself. Equivalently, if $V$ is a left module over $A$, then $\text{Hom}_k(V, W)$ becomes a right module over $A$ in this way, and if $V$ is a right module over $A$, then $\text{Hom}_k(V, W)$ becomes a left module over $A$.

Now let $A$ be a Lie algebra over $k$, and let $\rho$ be a representation of $A$ on $V$. As before, $\rho_a$ is defined as a module homomorphism from $V$ into $W$ for every $a \in A$. Under these conditions, one can check that $a \mapsto -\rho_a$ defines a representation of $A$ as a Lie algebra on $\text{Hom}_k(V, W)$.

### 7.3 Multilinear mappings

Let $k$ be a commutative ring with a multiplicative identity element, and let $n$ be a positive integer. Also let $V_1, \ldots, V_n$ be $n$ modules over $k$, and let $W$ be another module over $k$. A mapping $\mu$ from $V_1 \times \cdots \times V_n$ into $W$ is said to be multilinear over $k$ if $\mu$ is linear over $k$ in each variable separately. This reduces to ordinary linearity over $k$ when $n = 1$, and to bilinearity over $k$ when $n = 2$.

The space of mappings $\mu$ from $\prod_{j=1}^n V_j$ into $W$ that are multilinear over $k$ may be denoted $L(V_1, \ldots, V_n; W)$, or $L_k(V_1, \ldots, V_n; W)$ to indicate the role of $k$. It is easy to see that $L_k(V_1, \ldots, V_n; W)$ is a module over $k$ with respect
to pointwise addition and scalar multiplication of mappings from \( \prod_{j=1}^{n} V_j \) into \( W \). More precisely, \( L_k(V_1, \ldots, V_n; W) \) may be considered as a submodule of the module of all \( W \)-valued functions on \( \prod_{j=1}^{n} V_j \). Of course, \( L_k(V_1, \ldots, V_n; W) \) reduces to \( \text{Hom}_k(V_1, W) \) when \( n = 1 \).

Suppose for the moment that \( V_1, \ldots, V_n \) are the same module \( V \) over \( k \), and that \( n \geq 2 \). Let us say that \( \mu \in L_k(V_1, \ldots, V_n; W) \) is symmetric if \( \mu(v_1, \ldots, v_n) \) is invariant under permutations of the variables \( v_1, \ldots, v_n \). This reduces to the earlier notion of symmetry for bilinear mappings when \( n = 2 \). Similarly, let us say that \( \mu \in L_k(V_1, \ldots, V_n; W) \) is antisymmetric if

(7.3.1) \[
\mu(v_1, \ldots, v_{j-1}, v_l, v_{j+1}, \ldots, v_{l-1}, v_l, v_{j+1}, \ldots, v_n) = -\mu(v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_{l-1}, v_l, v_{j+1}, \ldots, v_n)
\]

for every \( v_1, \ldots, v_n \in V \) and \( 1 \leq j < l \leq n \), which is to say that interchanging two of the variables corresponds to taking the additive inverse of the value of \( \mu \). This reduces to the earlier notion of antisymmetry for bilinear mappings when \( n = 2 \). As usual, it is sometimes better to ask that

(7.3.2) \[
\mu(v_1, \ldots, v_n) = 0
\]

whenever \( v_j = v_l \) for some \( j \neq l \). This implies that \( \mu \) is antisymmetric, by the same type of argument as for bilinear mappings. If \( 1 + 1 \) is invertible in \( k \), then this condition holds when \( \mu \) is antisymmetric, as before.

Let \( V_1, \ldots, V_n \) be \( n \) modules over \( k \) for some \( n \geq 1 \) again, and let \( j \) be an integer with \( 1 \leq j \leq n \). Also let \( A_j \) be a module homomorphism from \( V_j \) into itself, and let \( \mu \) be a mapping from \( \prod_{j=1}^{n} V_j \) into \( W \) that is multilinear over \( k \). If \( v_1 \in V_1, \ldots, v_n \in V_n \), then put

(7.3.3) \[
(\widetilde{A}_j(\mu))(v_1, \ldots, v_n) = \mu(v_1, \ldots, v_{j-1}, A_j(v_j), v_{j+1}, \ldots, v_n).
\]

This defines a mapping \( \widetilde{A}_j(\mu) \) from \( \prod_{j=1}^{n} V_j \) into \( W \), which corresponds to composing \( \mu \) with \( A_j \) in the \( j \)th variable. Note that \( \widetilde{A}_j(\mu) \) is multilinear over \( k \), because \( A_j \) is linear over \( k \). It is easy to see that \( \widetilde{A}_j \) is linear over \( k \), as a mapping from \( L(V_1, \ldots, V_n; W) \) into itself. The mapping

(7.3.4) \[
A_j \mapsto \widetilde{A}_j
\]

is linear over \( k \) as well, as a mapping from \( \text{Hom}_k(V_j, V_j) \) into the space of module homomorphisms from \( L(V_1, \ldots, V_n; W) \) into itself.

Let \( B_j \) be another module homomorphism from \( V_j \) into itself, and let \( \mu \) be a mapping from \( \prod_{j=1}^{n} V_j \) into \( W \) that is multilinear over \( k \) again. The composition \( B_j \circ A_j \) is \( A_j \) and \( B_j \) is a module homomorphism from \( V_j \) into itself too, so that \( (B_j \circ A_j)(\mu) \) can be defined as a multilinear mapping from \( \prod_{j=1}^{n} V_j \) into \( W \) as before. More precisely, if \( v_1 \in V_1, \ldots, v_n \in V_n \), then

(7.3.5) \[
((B_j \circ A_j)(\mu))(v_1, \ldots, v_n) = \mu(v_1, \ldots, v_{j-1}, (B_j \circ A_j)(v_j), v_{j+1}, \ldots, v_n) = \mu(v_1, \ldots, v_{j-1}, B_j(A_j(v_j)), v_{j+1}, \ldots, v_n).
\]
This is the same as
\[(7.3.6) \quad (\overline{B_j}(\mu))(v_1, \ldots, v_{j-1}, A_j(v_j), v_{j+1}, \ldots, v_n) = (\overline{A_j}(\overline{B_j}(\mu)))(v_1, \ldots, v_n),\]
so that
\[(7.3.7) \quad (B_j \circ A_j) = \overline{A_j} \circ \overline{B_j}\]
as mappings from \(L(V_1, \ldots, V_n; W)\) into itself. Of course, this was mentioned in Section 7.2 when \(n = 1\).

### 7.4 Boundedness and multilinearity

Let \(k\) be a field with an absolute value function \(|\cdot|\), let \(n\) be a positive integer, and let \(V_1, \ldots, V_n\) and \(W\) be vector spaces over \(k\). Also let \(N_{V_1}, \ldots, N_{V_n}\) and \(N_W\) be seminorms on \(V_1, \ldots, V_n\) and \(W\), respectively, and with respect to \(|\cdot|\) on \(k\). A multilinear mapping \(\mu\) from \(V_1 \times \cdots \times V_n\) into \(W\) is said to be bounded with respect to \(N_{V_1}, \ldots, N_{V_n}\) and \(N_W\) if there is a nonnegative real number \(C\) such that
\[(7.4.1) \quad N_W(\mu(v_1, \ldots, v_n)) \leq C N_{V_1}(v_1) \cdots N_{V_n}(v_n)\]
for every \(v_1 \in V_1, \ldots, v_n \in V_n\). This reduces to the earlier definitions of boundedness for linear and bilinear mappings when \(n = 1\) and \(n = 2\), respectively, as in Sections 1.9 and 1.13.

Let \(BL(V_1, \ldots, V_n; W)\) be the space of bounded multilinear mappings from \(\prod_{j=1}^n V_j\) into \(W\), with respect to \(N_{V_1}, \ldots, N_{V_n}\) and \(N_W\). It is easy to see that this is a linear subspace of the space of all multilinear mappings from \(\prod_{j=1}^n V_j\) into \(W\). If \(\mu \in BL(V_1, \ldots, V_n; W)\), then put
\[(7.4.2) \quad \|\mu\| = \|\mu\|_{V_1, \ldots, V_n; W} = \inf \{C \geq 0 : (7.4.1) \text{ holds}\},\]
where more precisely the infimum is taken over all nonnegative real numbers \(C\) such that (7.4.1) holds for every \(v_1 \in V_1, \ldots, v_n \in V_n\). This reduces to the operator seminorm of a bounded linear mapping when \(n = 1\), as in (1.9.3). As before, the infimum in (7.4.2) is automatically attained, so that (7.4.1) holds with \(C = \|\mu\|\). One can check that (7.4.2) defines a seminorm on \(BL(V_1, \ldots, V_n; W)\) with respect to \(|\cdot|\), and that (7.4.2) is a norm on \(BL(V_1, \ldots, V_n; W)\) when \(N_W\) is a norm on \(W\). Similarly, if \(N_W\) is a semi-ultranorm on \(W\), then (7.4.2) is a semi-ultranorm on \(BL(V_1, \ldots, V_n; W)\).

Suppose for the moment that \(n \geq 2\), and let \(\mu\) be a multilinear mapping from \(\prod_{j=1}^n V_j\) into \(W\). If \(v_n \in V_n\), then
\[(7.4.3) \quad \mu_{v_n}(v_1, \ldots, v_{n-1}) = \mu(v_1, \ldots, v_{n-1}, v_n)\]
defines a multilinear mapping from \(V_1 \times \cdots \times V_{n-1}\) into \(W\). In addition,
\[(7.4.4) \quad T_{\mu}(v_n) = \mu_{v_n}\]
7.4. BOUNDEDNESS AND MULTILINEARITY

defines a linear mapping from \( V_n \) into \( L(V_1, \ldots, V_{n-1}; W) \). Note that every linear mapping from \( V_n \) into \( L(V_1, \ldots, V_{n-1}; W) \) corresponds to a multilinear mapping from \( \prod_{j=1}^{n} V_j \) into \( W \) in this way. If \( \mu \) is also bounded as a multilinear mapping, then

\[
N_W(\mu_{v_n}(v_1, \ldots, v_{n-1})) = N_W(\mu(v_1, \ldots, v_n))
\]
(7.4.5)

\[
\leq \|\mu\|_{V_1,\ldots,V_{n-1};W} N_{V_1}(v_1) \cdots N_{V_{n-1}}(v_{n-1}) N_{V_n}(v_n)
\]

for every \( v_1 \in V_1, \ldots, v_{n-1} \in V_{n-1}, v_n \in V_n \). This implies that \( \mu_{v_n} \) is bounded as a multilinear mapping for each \( v_n \in V_n \), with

\[
\|\mu_{v_n}\|_{V_1,\ldots,V_{n-1};W} \leq \|\mu\|_{V_1,\ldots,V_{n-1};W} N_{V_n}(v_n).
\]
(7.4.6)

Using \( \|\cdot\|_{V_1,\ldots,V_{n-1};W} \) on \( BL(V_1, \ldots, V_{n-1}; W) \), we get that (7.4.4) is bounded as a linear mapping from \( V_n \) into \( BL(V_1, \ldots, V_{n-1}; W) \), with

\[
\|\mu_{v_n}\|_{V_1,\ldots,V_{n-1};W} \leq \|T_{\mu}\|_{op} N_{V_n}(v_n).
\]
(7.4.7)

Conversely, suppose that (7.4.4) is a bounded linear mapping from \( V_n \) into \( BL(V_1, \ldots, V_{n-1}; W) \). If \( v_n \in V_n \), then \( \mu_{v_n} \in BL(V_1, \ldots, V_{n-1}; W) \), with

\[
\|\mu_{v_n}\|_{V_1,\ldots,V_{n-1};W} \leq \|T_{\mu}\|_{op} N_{V_n}(v_n).
\]
(7.4.8)

If \( v_1 \in V_1, \ldots, v_{n-1} \in V_{n-1} \), then we get that

\[
N_W(\mu(v_1, \ldots, v_{n-1}, v_n)) = N_W(\mu_{v_n}(v_1, \ldots, v_{n-1}))
\]
(7.4.9)

\[
\leq \|\mu_{v_n}\|_{V_1,\ldots,V_{n-1};W} N_{V_1}(v_1) \cdots N_{V_{n-1}}(v_{n-1}) \leq \|T_{\mu}\|_{op} N_{V_1}(v_1) \cdots N_{V_{n-1}}(v_{n-1}) N_{V_n}(v_n).
\]

This implies that \( \mu \in BL(V_1, \ldots, V_n; W) \), with

\[
\|\mu\|_{V_1,\ldots,V_n;W} \leq \|T_{\mu}\|_{op}.
\]
(7.4.10)

It follows that

\[
\|T_{\mu}\|_{op} = \|\mu\|_{V_1,\ldots,V_n;W},
\]
(7.4.11)

by (7.4.7).

Suppose now that \( n \geq 1, 1 \leq j \leq n \), and that \( A_j \) is a bounded linear mapping from \( V_j \) into itself, with respect to \( N_{V_j} \) on \( V_j \). Let \( \mu \) be a bounded multilinear mapping from \( \prod_{j=1}^{n} V_j \) into \( W \), and let \( \bar{A}_j(\mu) \) be the multilinear mapping from \( \prod_{j=1}^{n} V_j \) into \( W \) corresponding to \( \mu \) and \( A_j \) as in the previous section. If \( v_1 \in V_1, \ldots, v_n \in V_n \), then

\[
N_W((\bar{A}_j(\mu))(v_1, \ldots, v_n)) = N_W(\mu(v_1, \ldots, v_{j-1}, A_j(v_j), v_{j+1}, v_n))
\]
(7.4.12)

\[
\leq \|\mu\|_{V_1,\ldots,V_n;W} N_{V_1}(v_1) \cdots N_{V_{j-1}}(v_{j-1}) N_{V_j}(A_j(v_j)) N_{V_{j+1}}(v_{j+1}) \cdots N_{V_n}(v_n)
\]

\[
\leq \|A_j\|_{op, V_j} \|\mu\|_{V_1,\ldots,V_n;W} N_{V_1}(v_1) \cdots N_{V_n}(v_n),
\]
where \( \| A_j \|_{\text{op}, V_j} \) is the usual operator seminorm of \( A_j \) associated to \( N_{V_j} \) on \( V_j \). This shows that \( \tilde{A}_j(\mu) \) is also bounded as a multilinear mapping, with
\[
(7.4.13) \quad \| \tilde{A}_j(\mu) \|_{V_1, \ldots, V_n, W} \leq \| A_j \|_{\text{op}, V_j} \| \mu \|_{V_1, \ldots, V_n, W}.
\]
Similarly, let \( A \) be a bounded linear mapping from \( W \) into itself, with respect to \( N_W \) on \( W \). It is easy to see that the composition \( A \circ \mu \) of \( \mu \) with \( A \) is multilinear over \( k \), as a mapping from \( \prod_{j=1}^{n} V_j \) into \( W \). If \( v_1 \in V_1, \ldots, v_n \in V_n \), then
\[
N_W(A(\mu(v_1, \ldots, v_n))) \leq \| A \|_{\text{op}, W} N_W(\mu(v_1, \ldots, v_n)) \leq \| A \|_{\text{op}, W} \| \mu \|_{V_1, \ldots, V_n, W} N_{V_1}(v_1) \cdots N_{V_n}(v_n),
\]
where \( \| A \|_{\text{op}, W} \) is the operator seminorm of \( A \) associated to \( N_W \) on \( W \). Thus \( A \circ \mu \) is bounded as a multilinear mapping too, with
\[
(7.4.15) \quad \| A \circ \mu \|_{V_1, \ldots, V_n, W} \leq \| A \|_{\text{op}, W} \| \mu \|_{V_1, \ldots, V_n, W}.
\]

### 7.5 Representations on multilinear mappings

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( V_1, \ldots, V_n \) and \( W \) be modules over \( k \). The space \( L_k(V_1, \ldots, V_n; W) \) of mappings from \( \prod_{j=1}^{n} V_j \) into \( W \) that are multilinear over \( k \) is a module over \( k \) with respect to pointwise addition and scalar multiplication of mappings, as in Section 7.3.

Let \( A \) be an associative algebra over \( k \), and suppose for the moment that \( W \) is a left module over \( A \). If \( a \in A \) and \( \mu \in L(V_1, \ldots, V_n; W) \), then
\[
(7.5.1) \quad a \cdot (\mu(v_1, \ldots, v_n))
\]
defines a multilinear mapping from \( \prod_{j=1}^{n} V_j \) into \( W \) as well. It is easy to see that this makes \( L(V_1, \ldots, V_n; W) \) into a left module over \( A \).

Suppose that \( V_j \) is a left module over \( A \) for some \( j, 1 \leq j \leq n \). If \( a \in A \) and
\[
\mu(v_1, \ldots, v_j-1, a \cdot v_j, v_{j+1}, \ldots, v_n)
\]
defines another multilinear mapping from \( \prod_{j=1}^{n} V_j \) into \( W \). This defines an opposite algebra homomorphism from \( A \) into the algebra of module homomorphisms from \( L(V_1, \ldots, V_n; W) \) into itself, as in Section 7.3. Equivalently, \( L(V_1, \ldots, v_n; W) \) may be considered as a right module over \( A \) in this way. This was mentioned in Section 7.2 when \( n = 1 \).

Suppose from now on in this section that \( A \) is a Lie algebra over \( k \). If \( W \) is a module over \( A \) as a Lie algebra over \( k \), \( a \in A \), and \( \mu \in L(V_1, \ldots, V_n; W) \), then \( (7.5.1) \) defines an element of \( L(V_1, \ldots, V_n; W) \), as before. This makes \( L(V_1, \ldots, V_n; W) \) into a module over \( A \) as a Lie algebra over \( k \). Similarly, if \( V_j \) is a module over \( A \) as a Lie algebra over \( k \) for some \( j, a \in A \), and \( \mu \) is in \( L(V_1, \ldots, V_n; W) \), then \( (7.5.2) \) is an element of \( L(V_1, \ldots, V_n; W) \), and hence
\[
(7.5.3) \quad -\mu(v_1, \ldots, v_j-1, a \cdot v_j, v_{j+1}, \ldots, v_n)
\]
is an element of $L(V_1, \ldots, V_n; W)$. One can check that $L(V_1, \ldots, V_n; W)$ is a module over $A$ as a Lie algebra over $k$ with respect to (7.5.3), which was mentioned in Section 7.2 when $n = 1$.

Suppose that $V_1, \ldots, V_n$ and $W$ are all modules over $A$ as a Lie algebra. Let $a \in A$ and $\mu \in L_k(V_1, \ldots, V_n; W)$ be given, and let us define $a \cdot \mu$ as a mapping from $\prod_{j=1}^n V_j$ into $W$, as follows. If $v_j \in V_j$ for each $j = 1, \ldots, n$, then we put

$$(a \cdot \mu)(v_1, \ldots, v_n) = a \cdot (\mu(v_1, \ldots, v_n))$$

(7.5.4)

$$+ \sum_{j=1}^n (-\mu(v_1, \ldots, v_{j-1}, a \cdot v_j, v_{j+1}, \ldots, v_n)).$$

It is easy to see that (7.5.4) is multilinear over $k$ as a mapping from $\prod_{j=1}^n V_j$ into $W$, and that (7.5.4) is linear in $a$ over $k$. One can verify that this makes $L(V_1, \ldots, V_n; W)$ into a module over $A$ as a Lie algebra over $k$, using the remarks in the preceding paragraph. More precisely, each term on the right side of (7.5.4) defines a Lie algebra representation of $A$ on $L(V_1, \ldots, V_n; W)$, as before. One can check directly that these $n + 1$ representations of $A$ on $L(V_1, \ldots, V_n; W)$ commute with each other. Hence their sum defines a Lie algebra representation of $A$ on $L(V_1, \ldots, V_n; W)$ too, as in Section 6.5.

Suppose that $V_1, \ldots, V_n$ are the same module $V$ over $k$, with the same Lie algebra representation of $A$. Let $a \in A$ and $\mu \in L(V_1, \ldots, V_n; W)$ be given again. If $\mu$ is a symmetric multilinear mapping, then (7.5.4) is symmetric as well. If $\mu$ is antisymmetric, then one can verify that (7.5.4) is antisymmetric too. If $\mu(v_1, \ldots, v_n) = 0$ whenever $v_j = v_l$ for some $j \neq l$, then one can check that (7.5.4) satisfies the same condition. More precisely, if $v_j = v_l$ for some $j \neq l$, then this uses the antisymmetry of $\mu$ for the two terms on the right side of (7.5.4) that involve $a \cdot v_j$ and $a \cdot v_l$. Otherwise, one can apply the hypothesis on $\mu$ directly to the other terms on the right side of (7.5.4).

### 7.6 Centralizers and invariant elements

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an algebra over $k$ in the strict sense, where multiplication of $a, b \in A$ is denoted $a \cdot b$. The centralizer of a set $E \subseteq A$ in $A$ is the set of $a \in A$ that commute with every $x \in E$, which is to say that

$$a x = x a$$

(7.6.1)

for every $x \in E$. This is a submodule of $A$ as a module over $k$, and a subalgebra of $A$ when $A$ is associative. The center of $A$ is the centralizer of $A$ in itself.

Now let $(A, [\cdot, \cdot])$ be a Lie algebra over $k$. The centralizer of a set $E \subseteq A$ in $A$ as a Lie algebra is given by

$$C_A(E) = \{a \in A : [a, x]_A = 0 \text{ for every } x \in E\}.$$
150 CHAPTER 7. REPRESENTATIONS AND MULTILINEAR MAPPINGS

It is easy to see that this is a Lie subalgebra of $A$, using the Jacobi identity. The center of $A$ as a Lie algebra is given by

$$Z(A) = C_A(A) = \{a \in A : [a,x]_A = 0 \text{ for every } x \in A\},$$

which is automatically an ideal in $A$.

Note that (7.6.2) is contained in the centralizer of $A$ as an algebra in the strict sense, and that (7.6.3) is contained in the center of $A$ as an algebra in the strict sense. If $1 + 1$ is invertible in $k$, then (7.6.2) is the same as the centralizer of $E$ in $A$ as an algebra in the strict sense, and (7.6.3) is the same as the center of $A$ as an algebra in the strict sense.

If $A$ is an associative algebra over $k$, then $A$ is a Lie algebra over $k$ with respect to the corresponding commutator bracket $[x,y] = xy - yx$. Of course, (7.6.1) is the same as saying that $[a,x] = 0$. In this case, the centralizer of $E \subseteq A$ in $A$ as an associative algebra is the same as the centralizer of $E$ in $A$ as a Lie algebra with respect to the commutator bracket. In particular, the center of $A$ as an associative algebra is the same as the center of $A$ as a Lie algebra.

Let $\rho$ be a representation of $A$ as a Lie algebra over $k$ on a module $V$ over $k$. An element $v$ of $V$ is said to be invariant under $\rho$ if

$$\rho_a(v) = 0$$

for every $a \in A$, as on p31 of [24]. As usual, $V$ may be considered as a module over $A$ as a Lie algebra over $k$, with $a \cdot v = \rho_a(v)$ for every $a \in A$ and $v \in V$. Thus (7.6.4) can be reexpressed as saying that

$$a \cdot v = 0$$

for every $a \in A$.

If every $v \in V$ is invariant under $\rho$, then $\rho$ is said to act trivially on $V$. It is easy to see that the collection of $v \in V$ that are invariant under $\rho$ is a submodule of $V$, as a module over $k$. This defines a subrepresentation of $\rho$ on $V$, on which $\rho$ acts trivially.

If $A_0$ is any Lie subalgebra of $A$, then the restriction of $\rho_a$ to $a \in A_0$ defines a representation of $A_0$ on $V$, as a Lie algebra over $k$. If $v$ is any element of $V$, then the collection of $a \in A$ such that (7.6.4) or equivalently (7.6.5) holds is a Lie subalgebra of $A$.

Remember that the adjoint representation of $A$ is a representation of $A$ as a Lie algebra on itself, as a module over $k$. The collection of elements of $A$ that are invariant under the adjoint representation is the same as the center of $A$ as a Lie algebra.

### 7.7 Invariant multilinear mappings

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. Also let $V$, $W$ be modules over $k$, and let $\rho^V$, $\rho^W$ be representations of $A$ as a Lie algebra over $k$ on $V$, $W$, respectively.
Remember that the space $\text{Hom}_k(V, W)$ of homomorphisms from $V$ into $W$ as modules over $k$ is a module over $k$ too, with respect to pointwise addition and scalar multiplication of mappings. If $a \in A$, then $\rho^W_a$ and $\rho^W_a$ are homomorphisms from $V$ and $W$ into themselves, as modules over $k$. If $\phi \in \text{Hom}_k(V, W)$, then it follows that $\phi \circ \rho^V_a$ and $\rho^W_a \circ \phi$ are homomorphisms from $V$ into $W$ as well, as modules over $k$. The mappings

\begin{equation}
(7.7.1) \quad \phi \mapsto \rho^W_a \circ \phi
\end{equation}

and

\begin{equation}
(7.7.2) \quad \phi \mapsto -\phi \circ \rho^V_a
\end{equation}

define homomorphisms from $\text{Hom}_k(V, W)$ into itself, as a module over $k$. These define representations of $A$ as a Lie algebra over $k$ on $\text{Hom}_k(V, W)$, as in Section 7.5. We also saw that

\begin{equation}
(7.7.3) \quad \rho_a(\phi) = \rho^W_a \circ \phi - \phi \circ \rho^V_a
\end{equation}

defines a representation of $A$ as a Lie algebra over $k$ on $\text{Hom}_k(V, W)$. Remember that $\phi \in \text{Hom}_k(V, W)$ is said to be invariant under $\rho$ when (7.7.3) is equal to 0 for every $a \in A$, as in the previous section. This happens exactly when $\phi$ intertwines the representations $\rho^V$, $\rho^W$ of $A$ on $V$, $W$, respectively, as in Section 6.7.

Let $V$, $W$ be modules over $k$ again, and let $\rho^V$ be a representasion of $A$ as a Lie algebra over $k$ on $V$. Remember that the space $L_k(V, V; W)$ of mappings from $V \times V$ into $W$ that are bilinear over $k$ is a module over $k$ with respect to pointwise addition and scalar multiplication. If $a \in A$ and $\beta \in L_k(V, V; W)$, then $\rho^V_a$ is a homomorphism from $V$ into itself, as a module over $k$, and

\begin{equation}
(7.7.4) \quad \beta(u, \rho^V_a(v)), \quad \beta(u, \rho^V_a(v))
\end{equation}

define elements of $L_k(V, V; W)$. The mappings

\begin{equation}
(7.7.5) \quad \beta(u, v) \mapsto -\beta(u, \rho^V_a(v))
\end{equation}

and

\begin{equation}
(7.7.6) \quad \beta(u, v) \mapsto -\beta(u, \rho^V_a(v))
\end{equation}

define homomorphisms from $L_k(V, V; W)$ into itself, as a module over $k$. These define representations of $A$ as a Lie algebra over $k$ on $L_k(V, V; W)$, as in Section 7.5, and

\begin{equation}
(7.7.7) \quad (\rho_a(\beta))(u, v) = -\beta(\rho^V_a(u), v) - \beta(u, \rho^V_a(v))
\end{equation}

defines a representation of $A$ as a Lie algebra over $k$ on $L_k(V, V; W)$ too. The condition that

\begin{equation}
(7.7.8) \quad \rho_a(\beta) = 0
\end{equation}

as an element of $L_k(V, V; W)$ is the same as saying that (7.7.7) is equal to 0 for every $u, v \in V$, which means that $\rho^V_a$ is antisymmetric on $V$ with respect to $\beta$. Thus $\beta$ is invariant under the representation (7.7.7) of $A$ on $L_k(V, V; W)$ exactly when $\rho^V_a$ is antisymmetric on $V$ with respect to $\beta$ for every $a \in A$. 


Let \( V_1, \ldots, V_n \) and \( W \) be modules over \( k \), and let \( \mu \) be a mapping from \( \prod_{j=1}^n V_j \) into \( W \) that is multilinear over \( k \). Also let \( l \in \{1, \ldots, n\} \) be given, and let \( V_l^\mu \) be the set of \( v_l \in V_j \) such that
\[
(7.7.9) \quad \mu(v_1, \ldots, v_{l-1}, v_l, v_{l+1}, \ldots, v_n) = 0
\]
for every \( v_j \in V_j \) with \( 1 \leq j \leq n \) and \( j \neq l \). Note that \( V_l^\mu \) is a submodule of \( V_l \), as a module over \( k \). Let \( A \) be a Lie algebra over \( k \), and suppose that \( V_1, \ldots, V_n \) and \( W \) are modules over \( A \). Thus, for each \( a \in A \), \( a \cdot \mu \) can be defined as a mapping from \( \prod_{j=1}^n V_j \) into \( W \) that is multilinear over \( k \), as in Section 7.5. If \( a \cdot \mu = 0 \) as a mapping on \( \prod_{j=1}^n V_j \), and if \( v_l \in V_l^\mu \), then it is easy to see that \( a \cdot v_l \in V_l^\mu \) too. This means that \( V_l^\mu \) is a submodule of \( V_l \), as a module over \( A \), when \( a \cdot \mu = 0 \) for every \( a \in A \).

Let \( V \) be a module over \( k \), and let \( \beta \) be a mapping from \( V \times V \) into \( V \) that is bilinear over \( k \). Thus \( V \) is an algebra in the strict sense over \( k \), with respect to \( \beta \). Let \( A \) be a Lie algebra over \( k \), and suppose that \( V \) is a module over \( A \). If \( a \in A \), then \( a \cdot \beta \) is defined as a mapping from \( V \times V \) into \( V \) that is bilinear over \( k \) by
\[
(7.7.10) \quad (a \cdot \beta)(v, w) = a \cdot (\beta(v, w)) - \beta(a \cdot v, w) - \beta(v, a \cdot w)
\]
for every \( v, w \in V \), as in Section 7.5. Observe that \( a \cdot \beta = 0 \) as a mapping on \( V \times V \) exactly when \( \delta_a(v) = a \cdot v \) defines a derivation on \( V \) with respect to \( \beta \).

### 7.8 Traces of linear mappings

Let \( k \) be a commutative ring with a multiplicative identity element, let \( A_0 \) be a commutative associative algebra over \( k \), and let \( n \) be a positive integer. Remember that the space \( M_n(A_0) \) of \( n \times n \) matrices with entries in \( A_0 \) is an associative algebra over \( k \), using entrywise addition and scalar multiplication, and matrix multiplication. The trace of an element of \( M_n(A_0) \) defines a homomorphism from \( M_n(A_0) \) into \( A_0 \) as modules over \( k \), which satisfies
\[
(7.8.1) \quad \text{tr}(ab) = \text{tr}(ba)
\]
for every \( a, b \in M_n(A_0) \). Put
\[
(7.8.2) \quad B_0(a, b) = \text{tr}(ab)
\]
for every \( a, b \in M_n(A_0) \), which defines a mapping from \( M_n(A_0) \times M_n(A_0) \) into \( A_0 \). This mapping is bilinear over \( k \), and symmetric in \( a, b \).

Let \( a, b, x \in M_n(A_0) \) be given, and observe that
\[
(7.8.3) \quad B_0(a, b) = \text{tr}((a x) b) = \text{tr}(a (x b)) = B_0(a, x b).
\]
We also have that
\[
(7.8.4) \quad B_0(x a, b) = \text{tr}((x a) b) = \text{tr}(a (b x)) = B_0(a, b x).
\]
7.8. TRACES OF LINEAR MAPPINGS

Of course,
\[ a \mapsto ax, \quad a \mapsto xa \]
(7.8.5)
deﬁne homomorphisms from \( M_n(A_0) \) into itself, as a module over \( k \). Similarly,
\[ C_x(a) = [x, a] = xa - ax \]
(7.8.6)
defines \( C_x \) as a homomorphism from \( M_n(A_0) \) into itself, as a module over \( k \). It is easy to see that
\[ B_0(C_x(a), b) = -B_0(a, C_x(b)), \]
(7.8.7)
using (7.8.3) and (7.8.4).

Let \( k^n \) be the space of \( n \)-tuples of elements of \( k \), which is a (free) module over \( k \) with respect to coordinatewise addition and scalar multiplication. If \( a = (a_{j,l}) \) is an \( n \times n \) matrix with entries in \( k \), then
\[ (T_a(v))_j = \sum_{l=1}^{n} a_{j,l} v_l \]
(7.8.8)
defines a module homomorphism from \( k^n \) into itself, as usual. The mapping \( a \mapsto T_a \) defines an isomorphism from \( M_n(k) \) onto \( \text{Hom}_k(k^n, k^n) \), as associative algebras over \( k \). The trace of \( T_a \) is deﬁned as an element of \( k \) to be the trace of \( a \), which deﬁnes the trace as a homomorphism from \( \text{Hom}_k(k^n, k^n) \) into \( k \), as modules over \( k \). If \( R, T \in \text{Hom}_k(k^n, k^n) \), then we have that
\[ \text{tr}(R \circ T) = \text{tr}(T \circ R), \]
(7.8.9)
as in (7.8.1).

Let \( V \) be a module over \( k \) that is isomorphic to \( k^n \) as a module over \( k \), so that \( \text{Hom}_k(V, V) \) is isomorphic to \( \text{Hom}_k(k^n, k^n) \) as associative algebras over \( k \). The trace can be deﬁned as a homomorphism from \( \text{Hom}_k(V, V) \) into \( k \), as modules over \( k \), as before. One can check that this deﬁnition of the trace does not depend on the module isomorphism between \( V \) and \( k^n \), because of (7.8.9). Of course, if \( k \) is a ﬁeld, then an \( n \)-dimensional vector space \( V \) over \( k \) is isomorphic to \( k^n \) as a vector space over \( k \). This corresponds to choosing a basis for \( V \), and the trace of a linear mapping from \( V \) into itself does not depend on the choice of the basis.

Put
\[ B(T_1, T_2) = \text{tr}(T_1 \circ T_2) \]
(7.8.10)
for every \( T_1, T_2 \in \text{Hom}_k(V, V) \), which deﬁnes a symmetric bilinear form on \( \text{Hom}_k(V, V) \). If \( R \in \text{Hom}_k(V, V) \), then
\[ C_R(T) = [R, T] = R \circ T - T \circ R \]
(7.8.11)
deﬁnes a homomorphism from \( \text{Hom}_k(V, V) \) into itself, as a module over \( k \). As in (7.8.7), we have that
\[ B(C_R(T_1), T_2) = -B(T_1, C_R(T_2)) \]
for every \( T_1, T_2 \in \text{Hom}_k(V, V) \), so that \( C_R \) is antisymmetric on \( \text{Hom}_k(V, V) \) with respect to (7.8.10).
7.9 The Killing form

Let \( k \) be a commutative ring with a multiplicative identity element, and let \((A, [\cdot, \cdot]_A)\) be a Lie algebra over \( k \). Also let \( V \) be a module over \( k \), and let \( \rho \) be a representation of \( A \) as a Lie algebra on \( k \). Suppose that \( V \) is isomorphic to \( k^n \) as a module over \( k \) for some positive integer \( n \). If \( T \) is a homomorphism from \( V \) into itself, as a module over \( k \), then the trace \( \text{tr} T = \text{tr}_V T \) of \( T \) can be defined as an element of \( k \) as in the previous section. Put

\[
B_\rho(x, y) = \text{tr}_V(\rho_x \circ \rho_y)
\]

for every \( x, y \in A \), which is the trace of \( \rho_x \circ \rho_y \) as a module homomorphism from \( V \) into itself. This defines a mapping from \( A \times A \) into \( k \) that is bilinear over \( k \). Note that (7.9.1) is symmetric in \( x \) and \( y \).

Let \( w, x, y \in A \) be given, and observe that

\[
B_\rho([w, x]_A, y) = \text{tr}_V([\rho_w, \rho_x]_A \circ \rho_y) = \text{tr}_V([\rho_w, \rho_x] \circ \rho_y).
\]

The right side is equal to

\[
-\text{tr}_V(\rho_x \circ [\rho_w, \rho_y]),
\]

as in the previous section. It follows that

\[
B_\rho([w, x]_A, y) = B_\rho(x, [w, y]_A).
\]

If \( x \in A \), then \( \text{ad} x = \text{ad}_x \) is defined as a module homomorphism from \( A \) into itself by

\[
\text{ad}_x(z) = [x, z]_A
\]

for every \( z \in A \), as in Section 2.4. Thus (7.9.4) can be reformulated as saying that

\[
B_\rho(\text{ad}_w(x), y) = B_\rho(\text{ad}_w(y), x)
\]

for every \( w, x, y \in A \).

Remember that the space of bilinear forms on \( A \) may be considered as a module over \( A \), with respect to the adjoint representation on \( A \) and the trivial representation of \( A \) on \( k \), as in Section 7.5. Using this, (7.9.6) is the same as saying that (7.9.1) is invariant under this action of \( A \) on bilinear forms on \( A \), as in Section 7.7. This corresponds to Proposition 1.1 on p32 of [24], which was formulated for a field \( k \). Equivalently, (7.9.4) means that (7.9.1) is associative as a bilinear mapping on \( A \times A \), as in Section 6.10. See also p27 of [13].

If \( A \) is isomorphic to \( k^n \) as a module over \( k \) for some positive integer \( n \), then we can take \( V = A \) and \( \rho_x = \text{ad}_x \) in the previous paragraphs. In this case, (7.9.1) becomes

\[
b(x, y) = \text{tr}_A(\text{ad}_x \circ \text{ad}_y)
\]

for \( x, y \in A \). This is known as the Killing form on \( A \), as on p21 of [13], and Deﬁnition 1.2 on p32 of [24].
7.10 Invariant subspaces and traces

Let $k$ be a commutative ring with a multiplicative identity element, and let $V$ be a module over $k$. Remember that the collection $\text{Hom}_k(V,V)$ of all homomorphisms from $V$ into itself, as a module over $k$, is an associative algebra over $k$ with respect to composition of mappings. Let $W$ be a submodule of $V$, and let $A_W$ be the collection of all $T \in \text{Hom}_k(V,V)$ such that

$$T(W) \subseteq W.$$  

(7.10.1)

Note that $A_W$ is a subalgebra of $\text{Hom}_k(V,V)$. If $T \in A_W$, then let $T_W$ be the restriction of $T$ to $W$, which defines a module homomorphism from $W$ into itself. Of course,

$$T \mapsto T_W$$

defines an algebra homomorphism from $A_W$ into the algebra $\text{Hom}_k(W,W)$ of all module homomorphisms from $W$ into itself. Let $q$ be the canonical quotient mapping from $V$ onto the quotient module $V/W$. If $T \in A_W$, then there is a unique module homomorphism $T_{V/W}$ from $V/W$ into itself such that

$$T_{V/W} \circ q = q \circ T.$$  

(7.10.3)

It is easy to see that

$$T \mapsto T_{V/W}$$

defines an algebra homomorphism from $A_W$ into the algebra $\text{Hom}_k(V/W,V/W)$ of all module homomorphisms from $V/W$ into itself.

Suppose that $V/W$ is isomorphic to $k^n$ as a module over $k$ for some positive integer $n$, so that $V/W$ is a free module over $k$ of rank $n$. This means that there are $n$ elements $z_1, \ldots, z_n$ of $V$ such that every element of $V/W$ can be expressed in a unique way as a linear combination of $q(z_1), \ldots, q(z_n)$ with coefficients in $k$. Let $Z$ be the submodule of $V$ consisting of linear combinations of $z_1, \ldots, z_n$ with coefficients in $k$. More precisely, every element of $Z$ can be expressed in a unique way as a linear combination of $z_1, \ldots, z_n$ with coefficients in $k$, because of the analogous property of $q(z_1), \ldots, q(z_n)$ in $V/W$. Note that the restriction of $q$ to $Z$ defines an isomorphism from $Z$ onto $V/W$, as modules over $k$. It is easy to see that every element of $V$ can be expressed in a unique way as the sum of elements of $W$ and $Z$, so that $V$ may be identified with the direct sum of $W$ and $Z$, as a module over $k$. Suppose that $W$ is isomorphic to $k^m$ as a module over $k$ for some positive integer $m$ as well. Under these conditions, $V$ is isomorphic to $k^{m+n}$ as a module over $k$.

If $T \in A_W$, then the traces of $T$, $T_W$, and $T_{V/W}$ on $V$, $W$, and $V/W$, respectively, can be defined as elements of $k$, as in Section 7.8. Observe that

$$\text{tr}_V T = \text{tr}_W T_W + \text{tr}_{V/W} T_{V/W},$$

where the subscripts indicate the spaces on which the traces are taken. In particular, if

$$T(V) \subseteq W,$$  

(7.10.6)
then \( T_{V/W} = 0 \), and

\[
(7.10.7) \quad \text{tr}_V T = \text{tr}_W T_V.
\]

If \( T_1, T_2 \in \text{Hom}_k(V, V) \), then put

\[
(7.10.8) \quad B_V(T_1, T_2) = \text{tr}_V(T_1 \circ T_2),
\]

as in Section 7.8. Let \( B_W(\cdot, \cdot) \) and \( B_{V/W}(\cdot, \cdot) \) be the analogous bilinear forms on \( \text{Hom}_k(W, W) \) and \( \text{Hom}_k(V/W, V/W) \), respectively. Suppose that \( T_1, T_2 \in \mathcal{A}_W \), and let \( T_1, T_2, T_{1, V/W}, T_{2, V/W} \in \text{Hom}_k(V, W) \) and \( T_{1, V/W}, T_{2, V/W} \in \text{Hom}_k(V/W, V/W) \) be as before. Note that

\[
(7.10.9) \quad (T_1 \circ T_2)_W = T_1,_{V/W} \circ T_2,_{V/W}, \quad (T_1 \circ T_2)_{V/W} = T_1,_{V/W} \circ T_2,_{V/W},
\]

because (7.10.2) and (7.10.4) are algebra homomorphisms. It follows that

\[
(7.10.10) \quad B_V(T_1, T_2) = B_W(T_1, T_2) + B_{V/W}(T_{1, V/W}, T_{2, V/W}),
\]

by applying (7.10.5) to \( T = T_1 \circ T_2 \).

Let \((A;[,\cdot])\) be a Lie algebra over \( k \), and let \( \text{ad}_{A,x} \) be the restriction of \( \text{ad}_{x} \) to \( A \), which maps \( B \) into itself, because \( B \) is an ideal in \( A \). Similarly, let \( \text{ad}_{A,B} \) be the mapping from \( A/B \) into itself which is induced by \( \text{ad}_{x} \) on \( A \). Suppose that \( B \) and \( A/B \) are isomorphic as modules over \( k \) to \( k^m \) and \( k^n \), respectively, for some positive integers \( m \) and \( n \). This implies that \( A \) is isomorphic to \( k^{m+n} \) as a module over \( k \), as before. If \( x, y \in A \), then

\[
(7.10.11) \quad \text{ad}_{x}(z) = \text{ad}_{A,x}(z) = [x, z]
\]

for every \( z \in A \), as in Section 2.4. Let \( (\text{ad}_{x})_B \) be the restriction of \( \text{ad}_{x} \) to \( B \), which maps \( B \) into itself, because \( B \) is an ideal in \( A \). Similarly, let \( (\text{ad}_{x})_{A/B} \) be the mapping from \( A/B \) into itself which is induced by \( \text{ad}_{x} \) on \( A \). Suppose that \( B \) and \( A/B \) are isomorphic as modules over \( k \) to \( k^m \) and \( k^n \), respectively, for some positive integers \( m \) and \( n \). This implies that \( A \) is isomorphic to \( k^{m+n} \) as a module over \( k \), as before. If \( x, y \in A \), then

\[
(7.10.12) \quad \text{tr}_A(\text{ad}_{x} \circ \text{ad}_{y}) = \text{tr}_B((\text{ad}_{x})_B \circ (\text{ad}_{y})_B) + \text{tr}_{A/B}((\text{ad}_{x})_{A/B} \circ (\text{ad}_{y})_{A/B}),
\]

as in (7.10.10).

### 7.11 Radicals of bilinear mappings

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( V, W \) be modules over \( k \). Also let \( \beta \) be a mapping from \( V \times V \) into \( W \) that is bilinear over \( k \). Note that

\[
(7.11.1) \quad V^\beta = \{ u \in V : \beta(u, v) = 0 \text{ for every } v \in V \}
\]

is a submodule of \( V \). This may be called the radical of \( \beta \) in \( V \), as on p22 of [13].

Let \((A;[,\cdot])\) be a Lie algebra over \( k \), and let \( \beta \) be a mapping from \( A \times A \) into \( W \) that is bilinear over \( k \). Thus we take \( V = A \) in the preceding paragraph,
as a module over $k$. If $x \in A$, then $\text{ad}_x = \text{ad}_{A,x}$ is the module homomorphism from $A$ into itself defined by (7.10.11), as before. Suppose that

\begin{equation}
\beta(\text{ad}_w(x), y) = -\beta(x, \text{ad}_w(y))
\end{equation}

for every $w, x, y \in A$, which is the same as saying that

\begin{equation}
\beta([x, w], y) = \beta(x, [w, y])
\end{equation}

for every $w, x, y \in A$, as in Section 6.10. In this case, we may say that $\beta$ is associative as a bilinear form on $A$, as in Section 6.10. Equivalently, $\beta$ is invariant with respect to the representation on the space of bilinear mappings from $A \times A$ into $W$ corresponding to the adjoint representation on $A$ and the trivial representation on $W$, as in Section 7.7. It is easy that the radical $A^\beta$ of $\beta$ in $A$ is an ideal in $A$ as a Lie algebra over $k$, as on p22 of [13], and p44 of [24]. This may be considered as a particular case of statements in Sections 6.10 and 7.7.

Suppose that $A$ is isomorphic to $k^r$ as a module over $k$ for some positive integer $r$. If $x, y \in A$, then

\begin{equation}
\beta(x, y) = \text{tr}_A(\text{ad}_x \circ \text{ad}_y)
\end{equation}

is defined as an element of $k$, as in Section 7.8. This defines a symmetric bilinear mapping from $A \times A$ into $k$ that satisfies (7.11.2), as in Section 7.9. Thus the radical $A^\beta$ of (7.11.4) in $A$ is an ideal in $A$ as a Lie algebra over $k$, as in the preceding paragraph.

Let $B$ be an ideal in $A$ as a Lie algebra over $k$, so that the quotient $A/B$ is a Lie algebra over $k$ as well. Suppose that $B$ and $A/B$ are isomorphic to $k^m$ and $k^n$, respectively, as modules over $k$ for some positive integers $m$ and $n$. This implies that $A$ is isomorphic to $k^{m+n}$ as a module over $k$, as in the previous section. If $x, y \in A$, then let $(\text{ad}_x)_B$, $(\text{ad}_y)_B$ be the restrictions of $\text{ad}_x$, $\text{ad}_y$ to $B$, and let $(\text{ad}_x)_{A/B}$, $(\text{ad}_y)_{A/B}$ be the mappings from $A/B$ into itself induced by $\text{ad}_x$, $\text{ad}_y$, as before. If $x \in B$, then $\text{ad}_x$ maps $A$ into $B$, so that the induced mapping $(\text{ad}_x)_{A/B}$ is equal to 0. This implies that the second term on the right side of (7.10.12) is equal to 0 for every $y \in A$. It follows that

\begin{equation}
\text{tr}_A(\text{ad}_x \circ \text{ad}_y) = \text{tr}_B((\text{ad}_x)_B \circ (\text{ad}_y)_B)
\end{equation}

for every $x \in B$ and $y \in A$.

If $x \in B$ and $B$ is commutative as a Lie algebra, then $(\text{ad}_x)_B$ is equal to 0. This implies that the right side of (7.11.5) is equal to 0 for every $y \in A$. Under these conditions, we get that

\begin{equation}
\beta(x, y) = \text{tr}_A(\text{ad}_x \circ \text{ad}_y) = 0
\end{equation}

for every $x \in B$ and $y \in A$. This means that

\begin{equation}
B \subseteq A^\beta
\end{equation}

in this situation.
7.12 Tensor products

Let $k$ be a commutative ring with a multiplicative identity element, let $n \geq 2$ be an integer, and let $V_1, \ldots, V_n$ be $n$ modules over $k$. The tensor product

$$\bigotimes_{j=1}^{n} V_j = V_1 \bigotimes \cdots \bigotimes V_n$$

of these modules over $k$ is a module over $k$ with the following two properties. First, the tensor product comes equipped with a mapping from $\prod_{j=1}^{n} V_j$ into $\bigotimes_{j=1}^{n} V_j$ that is multilinear over $k$. The image of $(v_1, \ldots, v_n) \in \prod_{j=1}^{n} V_j$ in $\bigotimes_{j=1}^{n} V_j$ under this mapping is often expressed as $v_1 \otimes \cdots \otimes v_n$. Second, let $W$ be any module over $k$, and let $\mu$ be any mapping from $\prod_{j=1}^{n} V_j$ into $W$ that is multilinear over $k$. Under these conditions, $\mu$ can be expressed in a unique way as the composition of the mapping from $\prod_{j=1}^{n} V_j$ into $\bigotimes_{j=1}^{n} V_j$ just mentioned with a homomorphism from $\bigotimes_{j=1}^{n} V_j$ into $W$, as modules over $k$. Equivalently, this means that there is a unique module homomorphism $\tilde{\mu}$ from $\bigotimes_{j=1}^{n} V_j$ into $W$ such that

$$\tilde{\mu}(v_1 \otimes \cdots \otimes v_n) = \mu(v_1, \ldots, v_n)$$

for every $(v_1, \ldots, v_n) \in \prod_{j=1}^{n} V_j$. The tensor product is unique up to a suitable isomorphic equivalence.

Note that $\bigotimes_{j=1}^{n} V_j$ is generated as a module over $k$ by the associated image of $\prod_{j=1}^{n} V_j$. This means that every element of $\bigotimes_{j=1}^{n} V_j$ can be expressed as a finite sum of terms of the form $v_1 \otimes \cdots \otimes v_n$, where $v_j \in V_j$ for each $j = 1, \ldots, n$. This is clear from the standard construction of the tensor product, and it can also be obtained from the uniqueness of the tensor product.

Let $W_1, \ldots, W_n$ another collection of $n$ modules over $k$, and suppose that $\phi_j$ is a homomorphism from $V_j$ into $W_j$ for each $j = 1, \ldots, n$, as modules over $k$. Consider the mapping from $\prod_{j=1}^{n} V_j$ into $\bigotimes_{j=1}^{n} W_j$ that sends $(v_1, \ldots, v_n)$ in $\prod_{j=1}^{n} V_j$ to

$$\phi_1(v_1) \otimes \cdots \otimes \phi_n(v_n).$$

It is easy to see that this mapping is multilinear over $k$. This leads to a unique module homomorphism $\phi$ from $\bigotimes_{j=1}^{n} V_j$ into $\bigotimes_{j=1}^{n} W_j$ such that

$$\phi(v_1 \otimes \cdots \otimes v_n)$$

is equal to (7.12.3) for every $(v_1, \ldots, v_n) \in \prod_{j=1}^{n} V_j$.

Let $Z_1, \ldots, Z_n$ be another collection of $n$ modules over $k$, and let $\psi_j$ be a homomorphism from $W_j$ into $Z_j$ for each $j = 1, \ldots, n$, as modules over $k$. This leads to a module homomorphism $\psi$ from $\bigotimes_{j=1}^{n} W_j$ into $\bigotimes_{j=1}^{n} Z_j$, as in the preceding paragraph. Note that $\psi_j \circ \phi_j$ is a module homomorphism from $V_j$ into $Z_j$ for each $j = 1, \ldots, n$. One can check that $\psi \circ \phi$ is the same as the module homomorphism from $\bigotimes_{j=1}^{n} V_j$ into $\bigotimes_{j=1}^{n} Z_j$ obtained from $\psi_j \circ \phi_j$, $1 \leq j \leq n$, as in the previous paragraph.
7.13. FUNCTIONS ON SETS

Let $A$ be an associative algebra over $k$, and suppose that $V_l$ is also a left or right module over $A$ for some $l \in \{1, \ldots, n\}$. One can define an action of $A$ on $\bigotimes_{j=1}^n V_j$ on the left or the right, as appropriate, so that $\bigotimes_{j=1}^n V_j$ becomes a left or right module over $A$ too. More precisely, if $a \in A$, then the corresponding module homomorphism from $V_l$ into itself leads to a module homomorphism from $\bigotimes_{j=1}^n V_j$ into itself as before, using the identity mapping on $V_j$ when $j \neq l$.

Similarly, let $A$ be a Lie algebra over $k$, and suppose that $V_l$ is a module over $A$ for some $l \in \{1, \ldots, n\}$. One can define an action of $A$ on $\bigotimes_{j=1}^n V_j$ in the same way as in the preceding paragraph, so that $\bigotimes_{j=1}^n V_j$ becomes a module over $A$.

Suppose now that $V_l$ is a module over $A$ as a Lie algebra for each $l = 1, \ldots, n$, which leads to an action of $A$ on $\bigotimes_{j=1}^n V_j$ for each $l = 1, \ldots, n$, as in the previous paragraph. It is easy to see that these actions commute with each other on $\bigotimes_{j=1}^n V_j$. It follows that $\bigotimes_{j=1}^n V_j$ is a module over $A$ with respect to the sum of these actions, as in Section 7.1.

7.13 Functions on sets

Let $k$ be a commutative ring with a multiplicative identity element, and let $W$ be a module over $k$. Also let $X$ be a nonempty set, and let $c(X, W)$ be the space of $W$-valued functions on $X$. It is easy to see that $c(X, W)$ is a module over $k$ with respect to pointwise addition and scalar multiplication of functions. This is the same as the direct product of the family of copies of $W$ indexed by $X$.

If $f$ is a $W$-valued function on $X$, then the support of $f$ on $X$ is defined to be the set of $x \in X$ such that $f(x) \neq 0$. Let $c_{00}(X, W)$ be the subset of $c(X, W)$ consisting of functions with finite support. This is a submodule of $c(X, W)$, as a module over $k$, which corresponds to the direct sum of the family of copies of $W$ indexed by $X$. Of course, $c_{00}(X, W)$ is the same as $c(X, W)$ when $X$ has only finitely many elements.

In particular, we can take $W = k$, considered as a module over itself. If $x \in X$, then let $\delta_x$ be the $k$-valued function on $X$ equal to 1 at $x$ and to 0 elsewhere. Every element of $c_{00}(X, k)$ can be expressed in a unique way as a linear combination of finitely many $\delta_x$’s with coefficients in $k$.

Let $Z$ be another module over $k$. If $\phi$ is a homomorphism from $c_{00}(X, k)$ into $Z$, as modules over $k$, then $f(x) = \phi(\delta_x)$ defines a mapping from $X$ into $Z$. It is easy to see that $\phi$ is uniquely determined by $f$, and that every $Z$-valued function $f$ on $X$ corresponds to a module homomorphism $\phi$ from $c_{00}(X, k)$ into $Z$ in this way. The mapping $\phi \mapsto f$ defines an isomorphism from the space $\text{Hom}_k(c_{00}(X, k), Z)$ of module homomorphisms from $c_{00}(X, k)$ into $Z$ onto $c(X, Z)$, as modules over $k$.

Similarly, let $\phi$ be a homomorphism from $c_{00}(X, W)$ into $Z$, as modules over $k$. If $x \in X$ and $w \in W$, then $\delta_x w \in c_{00}(X, W)$, so that

$$\phi_x(w) = \phi(\delta_x w)$$ (7.13.1)
defines an element of $Z$. This defines $\phi_x$ as a module homomorphism from $W$ into $Z$, so that $x \mapsto \phi_x$ is an element of $c(X, \text{Hom}_k(W,Z))$. One can check that $\phi$ is uniquely determined by $x \mapsto \phi_x$, and that every mapping from $X$ into $\text{Hom}_k(W,Z)$ corresponds to a module homomorphism $\phi$ from $c_{00}(X,W)$ into $Z$ in this way. This defines an isomorphism between the space $\text{Hom}_k(c_{00}(X,W), Z)$ of module homomorphisms from $c_{00}(X,W)$ into $Z$ and $c(X, \text{Hom}_k(W,Z))$, as modules over $k$.

If $f \in c(X,k)$ and $w \in W$, then $f(x)w$ defines a $W$-valued function on $X$. This defines a mapping from $c(X,k) \times W$ into $c(X,W)$ that is bilinear over $k$. The restriction of this mapping to $c_{00}(X,k) \times W$ maps into $c_{00}(X,W)$.

Let $\mu$ be a mapping from $c_{00}(X,k) \times W$ into $Z$ that is bilinear over $k$. If $x \in X$, then put

$$\mu_x(w) = \mu(\delta_x, w)$$

for every $w \in W$, which defines a module homomorphism from $W$ into $Z$. If $f \in c_{00}(X,W)$, then

$$\sum_{x \in X} \mu_x(f(x))$$

defines an element of $Z$, where all but finitely many terms in the sum are equal to 0. This defines a homomorphism from $c_{00}(X,W)$ into $Z$, as modules over $k$. One can use this to check that $c_{00}(X,W)$ satisfies the requirements of the tensor product of $c_{00}(X,k)$ and $W$, as modules over $k$.

Let $V$ be another module over $k$, and suppose that $\phi$ is a homomorphism from $V$ into $c(X,Z)$, as modules over $k$. If $x \in X$ and $v \in V$, then let $\phi_x(v)$ be the value of $\phi(v)$ at $x$, as a $Z$-valued function on $X$. This defines $\phi_x$ as a module homomorphism from $V$ into $Z$ for each $x \in X$, so that $x \mapsto \phi_x$ is an element of $c(X, \text{Hom}_k(V,Z))$. Clearly $\phi$ is uniquely determined by $x \mapsto \phi_x$, and every mapping from $X$ into $\text{Hom}_k(V,Z)$ corresponds to a module homomorphism $\phi$ from $V$ into $c(X,Z)$ in this way. This defines an isomorphism between the space $\text{Hom}_k(V, c(X,Z))$ of module homomorphisms from $V$ into $c(X,Z)$ and $c(X, \text{Hom}_k(V,Z))$, as modules over $k$.

### 7.14 Some remarks about sums, products

Let $k$ be a commutative ring with a multiplicative identity element, and let $W$ be a module over $k$. Of course, $k$ may be considered as a module over itself, and

$$(t, w) \mapsto t \cdot w$$

defines a mapping from $k \times W$ into $W$ that is bilinear over $k$. Using this mapping, one can check that $W$ satisfies the requirements of $k \otimes W$.

Let $I$ be a nonempty set, and let $V_j$ be a module over $k$ for every $j \in I$. Thus $\bigoplus_{j \in I} V_j$ is a module over $k$, and one can verify that

$$(\bigoplus_{j \in I} V_j) \otimes W = \bigoplus_{j \in I} (V_j \otimes W)$$
7.15. SOME NATURAL ISOMORPHISMS

in a natural way. Of course, there is an analogous statement for tensor products with a direct sum in any factor.

If $Z$ is a module over $k$, then

$$(7.14.3) \quad \Hom_k(k, Z) = Z$$

in a natural way, because a mapping from $k$ into $Z$ that is linear over $k$ corresponds to multiplication by an element of $Z$. If $I$ and $V_j$, $j \in I$, are as in the previous paragraph, then

$$(7.14.4) \quad \Hom_k\left(\bigoplus_{j \in I} V_j, Z\right) = \prod_{j \in I} \Hom_k(V_j, Z)$$

in a natural way. More precisely, any module homomorphism from $\bigoplus_{j \in I} V_j$ into $Z$ leads to a module homomorphism from $V_l$ into $Z$ for every $l \in I$, using the natural inclusion of $V_l$ into $\bigoplus_{j \in I} V_j$. Conversely, if one has a module homomorphism from $V_l$ into $Z$ for every $l \in I$, then one can get a module homomorphism from $\bigoplus_{j \in I} V_j$ into $Z$ using the given homomorphisms on each coordinate of an element of $\bigoplus_{j \in I} V_j$, and adding the results in $Z$.

Similarly, let $V$ be a module over $k$, and let $Z_j$ be a module over $k$ for every $j \in I$. It is easy to see that

$$(7.14.5) \quad \Hom_k\left(V, \prod_{j \in I} Z_j\right) = \prod_{j \in I} \Hom_k(V, Z_j)$$

in a natural way.

## 7.15 Some natural isomorphisms

Let $k$ be a commutative ring with a multiplicative identity element, and let $V_1, \ldots, V_n$ be modules over $k$ for some integer $n \geq 2$. Also let $\sigma$ be a permutation on $\{1, \ldots, n\}$, which is to say a one-to-one mapping from $\{1, \ldots, n\}$ onto itself. Thus the tensor products $\bigotimes_{j=1}^n V_j$ and $\bigotimes_{j=1}^n V_{\sigma(j)}$ can be defined as modules over $k$ as in Section 7.12. There is a unique module homomorphism from $\bigotimes_{j=1}^n V_j$ into $\bigotimes_{j=1}^n V_{\sigma(j)}$ with

$$(7.15.1) \quad v_1 \otimes \cdots \otimes v_n \mapsto v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}$$

for every $(v_1, \ldots, v_n) \in \prod_{j=1}^n V_j$. More precisely, one can start with the mapping from $\prod_{j=1}^n V_j$ into $\bigotimes_{j=1}^n V_{\sigma(j)}$ defined by

$$(7.15.2) \quad (v_1, \ldots, v_n) \mapsto v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)},$$

which is multilinear over $k$.

Let $\tau$ be another permutation on $\{1, \ldots, n\}$, which leads to a unique module homomorphism from $\bigotimes_{j=1}^n V_{\sigma(j)}$ into $\bigotimes_{j=1}^n V_{\tau(\sigma(j))}$ with

$$(7.15.3) \quad v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)} \mapsto v_{\tau(\sigma(1))} \otimes \cdots \otimes v_{\tau(\sigma(n))}$$
for every \((v_1, \ldots, v_n) \in \prod_{j=1}^n V_j\), as before. The composition of this module homomorphism with the previous one that satisfies (7.15.1) is a module homomorphism from \(\bigotimes_{j=1}^n V_j\) into \(\bigotimes_{j=1}^n V_{\tau(\sigma(j))}\) with

\[
(7.15.4) \quad v_1 \otimes \cdots \otimes v_n \mapsto v_{\tau(\sigma(1))} \otimes \cdots \otimes v_{\tau(\sigma(n))}
\]

for every \((v_1, \ldots, v_n) \in \prod_{j=1}^n V_j\). This is the same as the homomorphism associated to \(\tau \circ \sigma\), by uniqueness. In particular, if \(\tau \circ \sigma\) is the identity mapping on \(\{1, \ldots, n\}\), then we get the identity mapping on \(\bigotimes_{j=1}^n V_j\). It follows that the module homomorphism from \(\bigotimes_{j=1}^n V_j\) into \(\bigotimes_{j=1}^n V_{\sigma(j)}\) that satisfies (7.15.1) is an isomorphism, whose inverse is the analogous homomorphism from \(\bigotimes_{j=1}^n V_{\sigma(j)}\) into \(\bigotimes_{j=1}^n V_j\) associated to the inverse of \(\sigma\).

Now let \(n_1\) and \(n_2\) be positive integers with \(n_1 + n_2 = n\). The tensor products \(\bigotimes_{j=1}^{n_1} V_j\) and \(\bigotimes_{l=1}^{n_2} V_{n_1+l}\) can be defined as before when \(n_1, n_2 \geq 2\), and otherwise they may be interpreted as being the given module over \(k\) when \(n_1\) or \(n_2\) is equal to 1. Thus the tensor product

\[
(7.15.5) \quad \left( \bigotimes_{j=1}^{n_1} V_j \right) \bigotimes \left( \bigotimes_{l=1}^{n_2} V_{n_1+l} \right)
\]

is defined as a module over \(k\) as well. It is well known that there is a natural isomorphism between \(\bigotimes_{j=1}^{n_1} V_j\) and (7.15.5), as modules over \(k\). More precisely, if \((v_1, \ldots, v_n) \in \prod_{j=1}^n V_j\), then \(v_1 \otimes \cdots \otimes v_n\) corresponds to

\[
(7.15.6) \quad (v_1 \otimes \cdots \otimes v_{n_1}) \otimes (v_{n_1+1} \otimes \cdots \otimes v_n)
\]

under this isomorphism.

Let \(V\) be a module over \(k\), and suppose that \(V_j = V\) for each \(j = 1, \ldots, n\). In this case,

\[
(7.15.7) \quad T^n V = \bigotimes_{j=1}^n V_j
\]

is called the \(n\)th tensor power of \(V\). We can interpret \(T^1 V\) as being equal to \(V\), as in the preceding paragraph. If \(\sigma\) is a permutation on \(\{1, \ldots, n\}\), then we get a module automorphism on \(T^n V\), as before. The elements of \(T^n V\) that are invariant under the module automorphisms associated to all permutations on \(\{1, \ldots, n\}\) are said to be symmetric, and form a submodule of \(T^n V\), as a module over \(k\).

If \(V_1, V_2, V_3\) are modules over \(k\), then \((V_1 \otimes V_2) \otimes V_3\) and \(V_1 \otimes (V_2 \otimes V_3)\) are both isomorphic to \(V_1 \otimes V_2 \otimes V_3\), as modules over \(k\), as before. This leads to a natural isomorphism between \((V_1 \otimes V_2) \otimes V_3\) and \(V_1 \otimes (V_2 \otimes V_3)\), as modules over \(k\). More precisely, if \(v_1 \in V_1\), \(v_2 \in V_2\), and \(v_3 \in V_3\), then \((v_1 \otimes v_2) \otimes v_3\) corresponds to \(v_1 \otimes (v_2 \otimes v_3)\) under this isomorphism.
Chapter 8

Formal series and ordered rings

8.1 Poles of finite order

Let $k$ be a commutative ring with a multiplicative identity element, let $A$ be a module over $k$, and let $T$ be an indeterminate. Consider the space $A((T))$ of formal series of the form

$$f(T) = \sum_{j=j_0}^{\infty} f_j T^j,$$

where $j_0 \in \mathbb{Z}$ and $f_j \in A$ for each $j \geq j_0$. More precisely, $A((T))$ may be defined as the space of $A$-valued functions on $\mathbb{Z}$ that are equal to 0 for all but finitely many negative integers. Thus (8.1.1) corresponds to $f \mapsto f_j$ as an $A$-valued function on $\mathbb{Z}$, where $f_j = 0$ when $j < j_0$. As in [4], elements of $A((T))$ may also be expressed as

$$f(T) = \sum_{j >> -\infty} f_j T^j,$$

where $f_j = 0$ for all but finitely many $j < 0$.

Note that $A((T))$ is a module over $k$ with respect to termwise addition and scalar multiplication of these formal series, which corresponds to pointwise addition and scalar multiplication of the associated $A$-valued functions on $\mathbb{Z}$. The space $A[[T]]$ of formal power series in $T$ with coefficients in $A$ can be identified with the submodule of $A((T))$ consisting of formal series $f(T)$ such that $f_j = 0$ for all $j < 0$. In particular, $A$ can be identified with the submodule of $A((T))$ consisting of formal series $f(T)$ with $f_j = 0$ when $j \neq 0$.

If $f(T) \in A((T))$ and $l \in \mathbb{Z}$, then

$$f(T) T^l = \sum_{j >> -\infty} f_j T^{j+l} = \sum_{j >> -\infty} f_{j-l} T^j.$$
defines an element of $A((T))$ as well. This is the same as $f(T)$ when $l = 0$, and agrees with the analogous definition on $A[[T]]$ in Section 4.3 when $l \geq 0$. In this situation, $f(T) \mapsto f(T) T^l$ is a module automorphism on $A((T))$ for every $l \in \mathbb{Z}$.

Let $n \in \mathbb{Z}$ be given, and let $(A[[T]])^n$ be the subset of $A((T))$ of formal series of the form $g(T) T^n$, where $g(T) \in A[[T]]$ is identified with an element of $A[[T]]$ as before. Equivalently, this is the set of $f(T) \in A((T))$ that can be expressed as in (8.1.1), with $j_0 \geq n$. If $n \geq 1$, then the elements of $(A[[T]])^n$ correspond to formal power series in $T$ with coefficients in $A$ that vanish to order $n - 1$, as in Section 4.3. Clearly $(A[[T]])^n$ is a submodule of $A((T))$ for every $n \in \mathbb{Z}$, with

$$A((T)) = \bigcup_{n=-\infty}^{\infty} (A[[T]])^n. \quad (8.1.4)$$

Let us say that the elements of a subset $E$ of $A((T))$ have **poles of bounded order** if $E \subseteq (A[[T]])^n$ for some $n \in \mathbb{Z}$. Of course, if $E$ has only finitely many elements, then the elements of $E$ have poles of bounded order. Suppose that $E$ is a submodule of $A((T))$, as a module over $k$. If $E$ is finitely generated, as a module over $k$, then it is easy to see that the elements of $E$ have poles of bounded order.

### 8.2 Sequences and series

Let us continue with the same notation and hypotheses as in the previous section. Let $l_0 \in \mathbb{Z}$ be given, and let

$$f_l(T) = \sum_{j \gg -\infty} f_{l,j} T^j \quad (8.2.1)$$

be an element of $A((T))$ for each integer $l \geq l_0$. As in Section 4.4, let us say that the sequence $\{f_l(T)\}_{l=l_0}^{\infty}$ is **termwise eventually constant** if for each $j \in \mathbb{Z}$ there is an integer $L_j \geq l_0$ such that $f_{l,j}$ does not depend on $l$ when $l \geq L_j$. Similarly, let us say that $\{f_l(T)\}_{l=l_0}^{\infty}$ **eventually agrees with** $f(T) \in A((T))$ **termwise** if for every $j \in \mathbb{Z}$ there is an integer $L \geq l_0$ such that

$$f_{l,j} = f_j \quad (8.2.2)$$

for every $l \geq L_j$. Of course, this implies that $\{f_l(T)\}_{l=l_0}^{\infty}$ is termwise eventually constant. If $\{f_l(T)\}_{l=l_0}^{\infty}$ is termwise eventually constant, and if the $f_l(T)$’s have poles of bounded order, then $\{f_l(T)\}_{l=l_0}^{\infty}$ eventually agrees with an element of $A((T))$ termwise. However, a sequence of elements of $A((T))$ may eventually agree termwise with an element of $A((T))$ without having poles of bounded order.
Let $\alpha \in k$ and $r \in \mathbb{Z}$ be given, as well as another sequence $\{g_i(T)\}_{i=0}^{\infty}$ of elements of $A((T))$. If $\{f_i(T)\}_{i=0}^{\infty}$ and $\{g_i(T)\}_{i=0}^{\infty}$ are termwise eventually constant, then so are $\{\alpha f_i(T)\}_{i=0}^{\infty}$, $\{f_i(T) + g_i(T)\}_{i=0}^{\infty}$, and $\{f_i(T) + g_i(T)\}_{i=0}^{\infty}$. If the $f_i(T)$'s and $g_i(T)$'s have poles of bounded order, then the $\alpha f_i(T)$'s and $f_i(T) + g_i(T)$'s have the same property. In this case, if $r_0 \in \mathbb{Z}$, then the $f_i(T)T^r$'s have poles of bounded order for $l \geq l_0$ and $r \geq r_0$. If $\{f_i(T)\}_{i=0}^{\infty}$ and $\{g_i(T)\}_{i=0}^{\infty}$ eventually agree with $f(T), g(T) \in A((T))$ termwise, respectively, then $\{\alpha f_i(T)\}_{i=0}^{\infty}$, $\{f_i(T) + g_i(T)\}_{i=0}^{\infty}$, and $\{f_i(T) + g_i(T)\}_{i=0}^{\infty}$ eventually agree with $\alpha f(T), f(T)T^r$, and $f(T) + g(T)$ termwise, respectively.

Let

\[(8.2.3) \quad a_l(T) = \sum_{j=-\infty}^{+\infty} a_{l+j} T^j \]

be an element of $A((T))$ for every integer $l \geq l_0$. Suppose that $\{a_l(T)\}_{l=0}^{\infty}$ eventually agrees with 0 termwise, and that the poles of the $a_l(T)$'s have bounded order. Under these conditions, the partial sums

\[(8.2.4) \quad \sum_{l=0}^{n} a_l(T) \]

are termwise eventually constant and have poles of bounded order. This implies that the partial sums $(8.2.4)$ eventually agree termwise with an element of $A((T))$, as before. Let us denote this element of $A((T))$ by

\[(8.2.5) \quad \sum_{l=0}^{\infty} a_l(T). \]

If $\alpha \in k$, then $\{\alpha a_l(T)\}_{l=0}^{\infty}$ eventually agrees with 0 termwise, the $\alpha a_l(T)$'s have poles of bounded order, and

\[(8.2.6) \quad \sum_{l=0}^{\infty} \alpha a_l(T) = \alpha \sum_{l=0}^{\infty} a_l(T). \]

Similarly, if $r \in \mathbb{Z}$, then $\{a_l(T)T^r\}_{l=0}^{\infty}$ eventually agrees with 0 termwise, the $a_l(T)T^r$'s have poles of bounded order (in $l$), and

\[(8.2.7) \quad \sum_{l=0}^{\infty} a_l(T)T^r = \left( \sum_{l=0}^{\infty} a_l(T) \right) T^r. \]

Let $\{b_l(T)\}_{l=0}^{\infty}$ be another sequence of elements of $A((T))$ that eventually agrees termwise with 0, and whose terms have poles of bounded order. This implies that $\{a_l(T) + b_l(T)\}_{l=0}^{\infty}$ eventually agrees termwise with 0 too, and that the $a_l(T) + b_l(T)$'s have poles of bounded order. It is easy to see that

\[(8.2.8) \quad \sum_{l=0}^{\infty} (a_l(T) + b_l(T)) = \sum_{l=0}^{\infty} a_l(T) + \sum_{l=0}^{\infty} b_l(T) \]

in this situation.
8.3 Formal series and module homomorphisms

Let $k$ be a commutative ring with a multiplicative identity element, let $A, B$ be modules over $k$, and let $T$ be an indeterminate. Thus $B((T))$ can be defined as a module over $k$ as in Section 8.1. Let $\phi$ be a homomorphism from $A$ into $B((T))$, as modules over $k$. If $a \in A$, then $\phi(a)$ can be expressed as

$$\phi(a) = \sum_{j=-\infty}^{\infty} \phi_j(a) T^j,$$

(8.3.1)

where $\phi_j(a) \in B$ for every $j \in \mathbb{Z}$, and $\phi_j(a) = 0$ for all but finitely many $j < 0$. More precisely, for each $j \in \mathbb{Z}$, $\phi_j$ is a homomorphism from $A$ into $B$, as modules over $k$. Conversely, let $\phi_j$ be a module homomorphism from $A$ into $B$ for every $j \in \mathbb{Z}$, and suppose that for each $a \in A$, $\phi_j(a) = 0$ for all but finitely many $j < 0$. Under these conditions, (8.3.1) defines an element of $B((T))$ for every $a \in A$, and this defines $\phi$ as a homomorphism from $A$ into $B((T))$, as modules over $k$.

Let $\phi$ be a module homomorphism from $A$ into $B((T))$ again. Let us say that $\phi$ has poles of bounded order if the set of $\phi(a)$ with $a \in A$ has poles of bounded order, as a subset of $B((T))$. Equivalently, this means that there is an integer $n(\phi)$ such that $\phi_j(a) = 0$ for every $a \in A$ and $j < n(\phi)$. This is the same as saying that $\phi_j = 0$ for all but finitely many $j < 0$, as homomorphisms from $A$ into $B$. If $A$ is finitely generated as a module over $k$, then this follows automatically from the fact that $\phi(a) \in B((T))$ for every $a \in A$.

Remember that the space $\text{Hom}_k(A, B)$ of module homomorphisms from $A$ into $B$ is a module over $k$ too, with respect to pointwise addition and scalar multiplication of mappings. Thus $(\text{Hom}_k(A, B))((T))$ can be defined as a module over $k$ as before. Let

$$\phi(T) = \sum_{l=l_0}^{\infty} \phi_l T^l$$

(8.3.2)

be an element of $(\text{Hom}_k(A, B))((T))$, so that $l_0 \in \mathbb{Z}$ and $\phi_l \in \text{Hom}_k(A, B)$ for every $l \geq l_0$. If $a \in A$, then

$$(\phi(T))(a) = \sum_{l=l_0}^{\infty} \phi_l(a) T^l$$

(8.3.3)

defines an element of $B((T))$, and the mapping from $a \in A$ to (8.3.3) is a homomorphism from $A$ into $B((T))$, as modules over $k$. This homomorphism has poles of finite order, and every module homomorphism from $A$ into $B((T))$ with poles of finite order corresponds to an element of $(\text{Hom}_k(A, B))((T))$ in this way.

The space $\text{Hom}_k(A, B((T)))$ of module homomorphisms from $A$ into $B((T))$ is a module over $k$ as well. It is easy to see that the collection of module homomorphisms from $A$ into $B((T))$ with poles of finite order is a submodule of $\text{Hom}_k(A, B((T)))$. The mapping from (8.3.2) to (8.3.3) defines an injective module homomorphism from $(\text{Hom}_k(A, B))((T))$ onto this submodule of
8.4. EXTENDING MODULE HOMOMORPHISMS

\text{Hom}_k(A, B((T)) \text{.}) If \( \phi(T) \in (\text{Hom}_k(A, B)((T)) \text{ and } r \in \mathbb{Z} \), then \( \phi(T) T^r \) defines an element of \( (\text{Hom}_k(A, B))((T)) \) too, as in Section 8.1. Note that for every \( a \in A \),
\begin{equation}
\phi(T) T^r(a) = (\phi(T))(a) T^r,
\end{equation}
as elements of \( B((T)) \).

8.4 Extending module homomorphisms

Let us continue with the notation and hypotheses in the previous section. Let \( \phi(T) \) be an element of \( (\text{Hom}_k(A, B))((T)) \) as in (8.3.2) again, and let

\begin{equation}
a(T) = \sum_{m=m_0}^{\infty} a_m T^m
\end{equation}
be an element of \( A((T)) \), where \( m_0 \in \mathbb{Z} \). Thus

\begin{equation}
(\phi(T))(a_m) = \sum_{l=l_0}^{\infty} \phi_l(a_m) T^l
\end{equation}
defines an element of \( B((T)) \) for every \( m \in \mathbb{Z} \), as in (8.3.3). Put

\begin{equation}
(\phi(T))(a(T)) = \sum_{m=m_0}^{\infty} (\phi(T))(a_m) T^m,
\end{equation}
where the series on the right can be defined as an element of \( B((T)) \) as in Section 8.2. This defines a homomorphism from \( A((T)) \) into \( B((T)) \), as modules over \( k \), associated to \( \phi(T) \).

It is easy to see that

\begin{equation}
(\phi(T))(a(T)) T^r = (\phi(T))(a(T)) T^r
\end{equation}
for every \( a(T) \in A((T)) \) and \( r \in \mathbb{Z} \). We also have that

\begin{equation}
(\phi(T))((A[[T]]) T^m) \subseteq (B[[T]]) T^{m+l_0}
\end{equation}
for every \( n \in \mathbb{Z} \). Note that (8.4.3) is linear over \( k \) in \( \phi(T) \). If \( r \in \mathbb{Z} \), then \( \phi(T) T^r \) defines an element of \( (\text{Hom}_k(A, B))((T)) \), as before. One can verify that

\begin{equation}
(\phi(T) T^r)(a(T)) = (\phi(T))(a(T)) T^r
\end{equation}
as elements of \( B((T)) \) for every \( a(T) \in A((T)) \).

Let \( a(T) \in A((T)) \) be given as in (8.4.1) again. As before, we take \( \phi_l = 0 \) when \( l < l_0 \), and \( a_m = 0 \) when \( m < m_0 \). Let \( n \in \mathbb{Z} \) be given, and observe that

\begin{equation}
\phi(a_{n-l}) = 0
\end{equation}
when \( l < l_0 \) and when \( n - l < m_0 \), which is to say that \( n - m_0 < l \). In particular, (8.4.7) holds for all but finitely many \( l \in \mathbb{Z} \), so that

\[
(8.4.8) \quad ((\phi(T))(a(T)))_n = \sum_{l=-\infty}^{\infty} \phi_l(a_{n-l})
\]
defines an element of \( B \). If \( n < l_0 + m_0 \), then (8.4.7) holds for every \( l \in \mathbb{Z} \), so that (8.4.8) is equal to 0. Consider

\[
(8.4.9) \quad (\phi(T))(a(T)) = \sum_{n=l_0+m_0}^{\infty} ((\phi(T))(a(T)))_n T^n
\]
as an element of \( B((T)) \). One can check that this is equivalent to (8.4.3).

Let \( C \) be a third module over \( k \), and let

\[
(8.4.10) \quad \psi(T) = \sum_{r=r_0}^{\infty} \psi_r T^r
\]
be an element of \( \text{Hom}_k(B,C)((T)) \), where \( r_0 \in \mathbb{Z} \). As usual, we take \( \psi_r = 0 \) when \( r < r_0 \). The composition \( \psi_r \circ \phi_l \) is defined as a module homomorphism from \( A \) into \( C \) for every \( l, r \in \mathbb{Z} \). Let \( n \in \mathbb{Z} \) be given, and note that

\[
(8.4.11) \quad \psi_r \circ \phi_{n-r} = 0
\]
when \( r < r_0 \) and when \( n - r < l_0 \), which means that \( n - l_0 < r \). It follows that (8.4.11) holds for all but finitely many \( r \in \mathbb{Z} \), so that

\[
(8.4.12) \quad (\psi(T) \circ \phi(T))_n = \sum_{r=-\infty}^{\infty} \psi_r \circ \phi_{n-r}
\]
defines a module homomorphism from \( A \) into \( C \). If \( n < l_0 + r_0 \), then (8.4.11) holds for every \( r \in \mathbb{Z} \), and (8.4.12) is equal to 0. Put

\[
(8.4.13) \quad \psi(T) \circ \phi(T) = \sum_{n=l_0+r_0}^{\infty} (\psi(T) \circ \phi(T))_n T^n,
\]
which defines an element of \( \text{Hom}_k(A,C)((T)) \). One can verify that the module homomorphism from \( A((T)) \) into \( C((T)) \) corresponding to (8.4.13) as before is the same as the composition of the homomorphisms from \( A((T)) \) into \( B((T)) \) and from \( B((T)) \) into \( C((T)) \) corresponding to \( \phi(T) \) and \( \psi(T) \), respectively.

### 8.5 Homomorphisms from \( A((T)) \) into \( B((T)) \)

Let \( k \) be a commutative ring with a multiplicative identity element again, let \( A, B \) be modules over \( k \), and let \( T \) be an indeterminate. Also let \( \phi \) be a homomorphism from \( A((T)) \) into \( B((T)) \), as modules over \( k \), and suppose that

\[
(8.5.1) \quad \phi(f(T))T = \phi(f(T)) T
\]
for every $f(T) \in A((T))$. This implies that

$$
(8.5.2) \quad \phi(f(T) T^r) = \phi(f(T)) T^r
$$

for every $f(T) \in A((T))$ and $r \in \mathbb{Z}$. Remember that $A[[T]]$ can be identified with the set of $f(T) \in A((T))$ such that $f_j = 0$ for every $j < 0$. Using (8.5.2), we get that $\phi$ is uniquely determined by its restriction to this subset of $A((T))$ corresponding to $A[[T]]$.

Suppose in addition that there is an integer $l_0(\phi)$ such that

$$
(8.5.3) \quad \phi(A[[T]]) T^{n} \subseteq (B[[T]]) T^{n+l_0(\phi)}
$$

for every $n \in \mathbb{Z}$. Note that this condition holds automatically when $\phi$ is obtained from an element of $(\text{Hom}_k(A, B))((T))$ as in the previous section. In order to verify this condition for any $\phi$ as in the preceding paragraph, it suffices to consider the case where $n = 0$, because of (8.5.2). The case where $n = 0$ can be reformulated as saying that if $f(T)$ corresponds to an element of $A[[T]]$, then $\phi(f(T)) T^{-l_0(\phi)}$ corresponds to an element of $B[[T]]$. This means that

$$
(8.5.4) \quad \tilde{\phi}(f(T)) = \phi(f(T)) T^{-l_0(\phi)}
$$

maps the subset of $A((T))$ corresponding to $A[[T]]$ into the subset of $B((T))$ corresponding to $B[[T]]$. Equivalently, the $n = 0$ case says that the collection of $\phi(f(T))$ with $f(T) \in A[[T]]$ has poles of bounded order in $B((T))$. Thus we may simply say that $\phi$ has poles of bounded order on $A[[T]]$ in this situation.

Under these conditions, one can check that $\phi$ is uniquely determined by its restriction to the subset of $A((T))$ corresponding to $A$. This can also be obtained from the analogous statement in Section 4.8, applied to the mapping from the subset of $A((T))$ corresponding to $A[[T]]$ into the subset of $B((T))$ corresponding to $B[[T]]$ given by (8.5.4). In this situation, the restriction of $\phi$ to the subset of $A((T))$ corresponding to $A$ has poles of bounded order, as in Section 8.3, and hence corresponds to an element of $(\text{Hom}_k(A, B))((T))$, as before. This element of $(\text{Hom}_k(A, B))((T))$ determines a homomorphism from $A((T))$ into $B((T))$, as in Section 8.4. In fact, $\phi$ is equal to this homomorphism on all of $A((T))$.

Let $\{a_l(T)\}_{l=0}^\infty$ be a sequence of elements of $A((T))$ starting at some $l_0 \in \mathbb{Z}$, and suppose that the set of $a_l(T)$’s, $l \geq l_0$, has poles of bounded order in $A((T))$. This implies that the set of $\phi(a_l(T))$’s, $l \geq l_0$, has poles of finite order in $B((T))$, by (8.5.3). Suppose that $\{a_l(T)\}_{l=0}^\infty$ also eventually agrees with some $a(T)$ in $A((T))$ termwise, as in Section 8.2. Under these conditions, one can check that $\{\phi(a_l(T))\}_{l=l_0}^\infty$ eventually agrees with $\phi(a(T))$ termwise. In particular, if $\{a_l(T)\}_{l=l_0}^\infty$ eventually agrees termwise with 0, then $\{\phi(a_l(T))\}_{l=l_0}^\infty$ eventually agrees termwise with 0. In this situation, $\sum_{l=l_0}^\infty a_l(T)$ and $\sum_{l=l_0}^\infty \phi(a_l(T))$ can be defined as elements of $A((T))$ and $B((T))$, respectively, as in Section 8.2. It is easy to see that

$$
(8.5.5) \quad \phi\left(\sum_{l=l_0}^\infty a_l(T)\right) = \sum_{l=l_0}^\infty \phi(a_l(T)),
$$

using the previous statement for the partial sums of these series.
8.6 Formal series and bilinear mappings

Let \( k \) be a commutative ring with a multiplicative identity element, let \( A, B, C \) be modules over \( k \), and let \( T \) be an indeterminate. Thus \( C((T)) \) can be defined as a module over \( k \) as in Section 8.1, and we let \( \beta \) be a mapping from \( A \times B \) into \( C((T)) \) that is bilinear over \( k \). If \( \alpha \in A \) and \( \beta \in B \), then \( \beta(a, b) \) can be expressed as

\[
(8.6.1) \quad \beta(a, b) = \sum_{r=-\infty}^{\infty} \beta_r(a, b) T^r,
\]

where \( \beta_r(a, b) \in C \) for every \( r \in \mathbb{Z} \), and

\[
(8.6.2) \quad \beta_r(a, b) = 0
\]

for all but finitely many \( r < 0 \). This defines \( \beta_r \) as a mapping from \( A \times B \) into \( C \) that is bilinear over \( k \) for every \( r \in \mathbb{Z} \). Conversely, if \( \beta_r \) is a mapping from \( A \times B \) into \( C \) that is bilinear over \( k \) for every \( r \in \mathbb{Z} \), and if for every \( a \in A \) and \( b \in B \) we have that (8.6.2) holds for all but finitely many \( r < 0 \), then (8.6.1) defines an element of \( C((T)) \) for every \( a \in A \) and \( b \in B \), and this defines a mapping from \( A \times B \) into \( C((T)) \) that is bilinear over \( k \).

Let \( \beta \) be a mapping from \( A \times B \) into \( C((T)) \) that is bilinear over \( k \) again. Let us say that \( \beta \) has poles of bounded order if the set of \( \beta(a, b) \) with \( a \in A \) and \( b \in B \) has poles of bounded order in \( C((T)) \). This means that there is an integer \( r(\beta) \) such that (8.6.2) holds for every \( a \in A, b \in B, \) and \( r < r(\beta) \). One can check that this holds automatically when \( A \) and \( B \) are finitely generated as modules over \( k \). If \( \beta \) has poles of bounded order, then \( \beta \) corresponds to a formal series in \( T \) with poles of finite order whose coefficients are bilinear mappings from \( A \times B \) into \( C \).

Let \( \beta \) be a mapping from \( A \times B \) into \( C((T)) \) that is bilinear over \( k \) and has poles of bounded order, so that there is an \( r(\beta) \in \mathbb{Z} \) such that (8.6.2) holds for every \( a \in A, b \in B, \) and \( r < r(\beta) \). Also let \( f(T) = \sum_{j=j_0}^{\infty} f_j T^j \in A((T)) \) and \( g(T) = \sum_{l=l_0}^{\infty} g_l T^l \in B((T)) \) be given, where \( j_0, l_0 \in \mathbb{Z} \). As before, we take \( f_j = 0 \) when \( j < j_0, \) and \( g_l = 0 \) when \( l < l_0, \) Let \( n \in \mathbb{Z} \) be given, and observe that

\[
(8.6.3) \quad \beta(f_j, g_{n-j}) = 0
\]

when \( j < j_0, \) and when \( n - j < l_0, \) which means that \( n - l_0 < j \). In particular, (8.6.3) holds for all but finitely many \( j \in \mathbb{Z} \), so that

\[
(8.6.4) \quad h_n = \sum_{j=-\infty}^{\infty} \beta(f_j, g_{n-j})
\]

defines an element of \( C((T)) \). Equivalently,

\[
(8.6.5) \quad \beta(f_{n-l}, g_l) = 0
\]
when \( l < l_0 \) and when \( n - l < j_0 \), which means that \( n - j_0 < l \). Thus (8.6.5) holds for all but finitely many \( l \in \mathbb{Z} \), and (8.6.4) is the same as

\[
\tag{8.6.6}
h_n = \sum_{l=-\infty}^{\infty} \beta(f_{n-l}, g_l).
\]

If \( n < j_0 + l_0 \), then (8.6.3) holds for every \( j \in \mathbb{Z} \), which is the same as saying that (8.6.5) holds for every \( l \in \mathbb{Z} \), so that \( h_n = 0 \). Note that the coefficient of \( T^r \) in \( h_n \) is equal to 0 when \( r < r(\beta) \), because of the corresponding hypothesis on \( \beta \). Put

\[
\tag{8.6.7}
h(T) = \sum_{n=j_0+l_0}^{\infty} h_n T^n,
\]

where the series on the right defines an element of \( C((T)) \) as in Section 8.2.

Put

\[
\tag{8.6.8}
\beta(f(T), g(T)) = h(T),
\]

which defines a mapping from \( A((T)) \times B((T)) \) into \( C((T)) \) that is bilinear over \( k \). This mapping agrees with the initial mapping from \( A \times B \) into \( C((T)) \), when \( A \) and \( B \) are identified with submodules of \( A((T)) \) and \( B((T)) \), respectively, as in Section 8.1. One can verify that

\[
\tag{8.6.9}
\beta(f(T) T^{m_1}, g(T) T^{m_2}) = \beta(f(T), g(T)) T^{m_1 + m_2}
\]

for every \( f(T) \in A((T)), g(T) \in B((T)) \), and \( m_1, m_2 \in \mathbb{Z} \). The coefficient of \( T^r \) in (8.6.7) is equal to 0 when

\[
\tag{8.6.10}
r < j_0 + l_0 + r(\beta),
\]

because of the analogous statement for \( h_n \). Equivalently, if \( f(T) \in (A[[T]]) T^{m_1} \) and \( g(T) \in (B[[T]]) T^{m_2} \) for some \( m_1, m_2 \in \mathbb{Z} \), then

\[
\tag{8.6.11}
\beta(f(T), g(T)) \in (C[[T]]) T^{m_1 + m_2 + r(\beta)}.
\]

Suppose that \( A = B \). If the initial mapping \( \beta \) from \( A \times A \) into \( C((T)) \) is symmetric or antisymmetric, then the extension of \( \beta \) to \( A((T)) \times A((T)) \) just defined has the same property, because (8.6.4) and (8.6.6) are the same. Similarly, if \( \beta(a, a) = 0 \) for every \( a \in A \), then

\[
\tag{8.6.12}
\beta(f(T), f(T)) = 0
\]

for every \( f(T) \in A((T)) \). To see this, it suffices to verify that

\[
\tag{8.6.13}
\sum_{j=-\infty}^{\infty} \beta(f_j, f_{n-j}) = 0
\]

for every \( n \in \mathbb{Z} \). Remember that \( \beta \) is antisymmetric on \( A \times A \) in this situation, as in Section 2.1. If \( n \) is odd, then (8.6.13) follows from the antisymmetry of \( \beta \) on \( A \times A \). If \( n \) is even, then (8.6.13) follows from the antisymmetry of \( \beta \) and the fact that \( \beta(f_{n/2}, f_{n/2}) = 0 \), by hypothesis.
8.7 Bilinear mappings on $A((T)) \times B((T))$

Let $k$ be a commutative ring with a multiplicative identity element again, let $A$, $B$, and $C$ be modules over $k$, and let $T$ be an indeterminate. Suppose that $\beta$ is a mapping from $A((T)) \times B((T))$ into $C((T))$ that is bilinear over $k$ and satisfies

$$\beta(f(T), g(T)) = \beta(f(T), g(T)) T$$

for every $f(T) \in A((T))$ and $g(T) \in B((T))$. This implies that

$$\beta(f(T) T^{m_1}, g(T) T^{m_2}) = \beta(f(T), g(T)) T^{m_1 + m_2}$$

for every $f(T) \in A((T))$, $g(T) \in B((T))$, and $m_1, m_2 \in \mathbb{Z}$. It is easy to see that $\beta$ is uniquely determined by its restriction to the subset of $A((T)) \times B((T))$ corresponding to $A[[T]] \times B[[T]]$, using (8.7.2).

As before, let us ask in addition that there be an integer $r(\beta)$ such that if $f(T) \in (A[[T]]) T^{m_1}$ and $g(T) \in (B[[T]]) T^{m_2}$ for some $m_1, m_2 \in \mathbb{Z}$, then

$$\beta(f(T), g(T)) \in (C[[T]]) T^{m_1 + m_2 + r(\beta)}.$$  

Remember that this condition holds when $\beta$ is obtained from a bilinear mapping from $A \times B$ into $C((T))$ with poles of bounded order as in the previous section. To verify this condition for any $\beta$ as in the preceding paragraph, it is enough to consider the case where $m_1 = m_2 = 0$, because of (8.7.2). This case can be reformulated as saying that if $f(T)$ and $g(T)$ correspond to elements of $A[[T]]$ and $B[[T]]$, respectively, then $\beta(f(T), g(T)) T^{-r(\beta)}$ corresponds to an element of $C[[T]]$. Equivalently, this means that

$$\tilde{\beta}(f(T), g(T)) = \beta(f(T), g(T)) T^{-r(\beta)}$$

maps the subset of $A((T)) \times B((T))$ corresponding to $A[[T]] \times B[[T]]$ into the subset of $C((T))$ corresponding to $C[[T]]$. The $m_1 = m_2 = 0$ case of (8.7.3) is the same as saying that the collection of $\beta(f(T), g(T))$ with $f(T) \in A[[T]]$ and $g(T) \in B[[T]]$ has poles of bounded order in $C((T))$. In this situation, we may simply say that $\beta$ has poles of bounded order on $A[[T]] \times B[[T]]$.

One can check that $\beta$ is uniquely determined by its restriction to the subset of $A((T)) \times B((T))$ that corresponds to $A \times B$ under these conditions. This can also be seen using the analogous statement in Section 4.8 for mappings from $A[[T]] \times B[[T]]$ into $C[[T]]$, applied to the mapping that corresponds to (8.7.4). Of course, the restriction of $\beta$ to $A \times B$ has poles of bounded order on $A \times B$. Thus the restriction of $\beta$ to $A \times B$ can be extended to $A((T)) \times B((T))$ as in the previous section. This extension agrees with $\beta$ on all of $A((T)) \times B((T))$ in this situation.

Let $\{a_m(T)\}_{m=m_0}^\infty$ and $\{b_r(T)\}_{r=r_0}^\infty$ be sequences of elements of $A((T))$ and $B((T))$, respectively. Suppose that the sets of $a_m(T)$’s, $m \geq m_0$, and $b_r(T)$’s, $r \geq r_0$, have poles of bounded order in $A((T))$ and $B((T))$, respectively. This implies that the set of $\beta(a_m(T), b_r(T))$’s, $m \geq m_0, r \geq r_0$, has poles of bounded
order in $C((T))$, by (8.7.3). Suppose that \( \{a_m(T)\}_{m=m_0}^{\infty} \) and \( \{b_r(T)\}_{r=r_0}^{\infty} \) eventually agree with some \( a(T) \in A((T)) \) and \( b(T) \in B((T)) \) termwise, respectively, as in Section 8.2. One can verify that \( \{\beta(a_r(T), b_r(T))\}_{r=\max(m_0, r_0)}^{\infty} \) eventually agrees termwise with \( \beta(a(T), b(T)) \).

Suppose now that \( \{a_m(T)\}_{m=m_0}^{\infty} \) and \( \{b_r(T)\}_{r=r_0}^{\infty} \) eventually agree termwise with 0 in \( A((T)) \) and \( B((T)) \), respectively, in addition to having poles of bounded order. This implies that \( \sum_{m=m_0}^{\infty} a_m(T) \) and \( \sum_{r=r_0}^{\infty} b_r(T) \) define elements of \( A((T)) \) and \( B((T)) \), respectively, as in Section 8.2. If \( N \) is an integer with \( N \geq m_0 + r_0 \), then put

\[
(8.7.5) \quad c_N(T) = \sum_{m=m_0}^{N-r_0} \beta(a_m(T), b_{N-m}(T)),
\]

which is an element of \( C((T)) \). Equivalently, this is the sum of

\[
(8.7.6) \quad \beta(a_m(T), b_r(T))
\]

over \( m \geq m_0 \) and \( r \geq r_0 \) with \( m + r = N \). Note that the set of \( c_N(T) \)'s, \( N \geq m_0 + r_0 \), has poles of bounded order in \( C((T)) \), because of the analogous statement for (8.7.6) in the preceding paragraph. One can check that (8.7.6) vanishes to arbitrarily large order in \( T \) when \( m \) or \( r \) is sufficiently large, because of (8.7.3). In particular, \( \{c_N(T)\}_{N=m_0+r_0}^{\infty} \) eventually agrees termwise with 0. Thus \( \sum_{N=m_0+r_0}^{\infty} c_N(T) \) defines an element of \( C((T)) \), as in Section 8.2. One can verify that

\[
(8.7.7) \quad \sum_{N=m_0+r_0}^{\infty} c_N(T) = \beta\left( \sum_{m=m_0}^{\infty} a_m(T), \sum_{r=r_0}^{\infty} b_r(T) \right),
\]

as in Section 4.2.

### 8.8 Algebras and modules over \( k((T)) \)

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( A \) be an algebra over \( k \) in the strict sense, where multiplication of \( a, b \in A \) is expressed as \( ab \). Also let \( T \) be an indeterminate, and let \( f(T) = \sum_{j=j_0}^{\infty} f_j T^j \) and \( g(T) = \sum_{l=l_0}^{\infty} g_l T^l \) be elements of \( A((T)) \), where \( j_0, l_0 \in \mathbb{Z} \). As usual, we take \( f_j = 0 \) when \( j < j_0 \), and \( g_l = 0 \) when \( l < l_0 \). Let \( n \in \mathbb{Z} \) be given, and observe that

\[
(8.8.1) \quad f_j g_{n-j} = 0
\]

when \( j < j_0 \) and when \( n - j < l_0 \), which means that \( n - l_0 < j \). In particular, (8.8.1) holds for all but finitely many \( j \in \mathbb{Z} \), so that

\[
(8.8.2) \quad h_n = \sum_{j=-\infty}^{\infty} f_j g_{n-j}
\]
reduces to a finite sum in $A$. Equivalently,

$$(8.8.3) \quad f_{n-l} g_l = 0$$

when $l < l_0$ and when $n - l < j_0$, which means that $n - j_0 < l$. Thus (8.8.3) holds for all but finitely many $l \in \mathbb{Z}$, and (8.8.2) is the same as

$$(8.8.4) \quad h_n = \sum_{l=-\infty}^{\infty} f_{n-l} g_l.$$ 

If $n < j_0 + l_0$, then (8.8.1) holds for every $j \in \mathbb{Z}$, which is the same as saying that (8.8.3) holds for every $l \in \mathbb{Z}$, so that $h_n = 0$. Put

$$(8.8.5) \quad f(T) g(T) = h(T) = \sum_{n=j_0+l_0}^{\infty} h_n T^n,$$

which defines another element of $A((T))$. Of course, this is the same as in (8.6.7), with the bilinear mapping $\beta$ given by multiplication on $A$. As before, one can check that

$$(8.8.6) \quad (f(T) T^{m_1}) (g(T) T^{m_2}) = (f(T) g(T)) T^{m_1+m_2}$$

for every $m_1, m_2 \in \mathbb{Z}$. If $f(T) \in (A[[T]]) T^{m_1}$ and $g(T) \in (A[[T]]) T^{m_2}$ for some $m_1, m_2 \in \mathbb{Z}$, then

$$(8.8.7) \quad f(T) g(T) \in (A[[T]]) T^{m_1+m_2}.$$ 

This extends multiplication on $A$ to a mapping from $A((T)) \times A((T))$ into $A((T))$ that is bilinear over $k$, which makes $A((T))$ into an algebra over $k$ in the strict sense. In particular, this agrees with the extension of multiplication on $A$ to $A[[T]]$ discussed in Section 4.6, so that $A[[T]]$ may be considered as a subalgebra of $A((T))$. If $A$ is a commutative algebra over $k$, then $A((T))$ is commutative as well, by the remark about symmetry of $\beta$ in the previous section. Similarly, if $A$ is an associative algebra over $k$, then one can verify that $A((T))$ is an associative algebra. If $A$ has a multiplicative identity element $e$, then $e$ corresponds to the multiplicative identity element in $A((T))$ too.

Suppose that $A$ is an associative algebra over $k$ with a multiplicative identity element $e$, and let $f(T) = \sum_{j=j_0}^{\infty} f_j T^j$ be an element of $A((T))$. Thus $f(T) T^{-j_0}$ corresponds to an element of $A[[T]]$. If $f_{j_0}$ is invertible in $A$, then $f(T) T^{-j_0}$ is invertible in $A[[T]]$, as in Section 4.7. This implies that $f(T)$ is invertible in $A((T))$.

Applying the earlier remarks to $k$ as a commutative associative algebra over itself, we get that $k((T))$ is a commutative associative algebra over $k$ with a multiplicative identity element. Let $A$ be a module over $k$, and let $f(T) \in k((T))$ and $g(T) \in A((T))$ be given. Under these conditions, $f(T) g(T)$ can be defined as an element of $A((T))$ as in (8.8.5), where the terms on the right side of (8.8.2) are defined using scalar multiplication on $A$. This is the same as extending scalar multiplication on $A$, as a mapping from $k \times A$ into $A$ that is bilinear over $k$, to
8.9. ABSOLUTE VALUES ON $K((T))$

A mapping from $k((T)) \times A((T))$ into $A((T))$, as in the previous section. One can check that this makes $A((T))$ into a module over $k((T))$. Note that (8.8.6) holds for every $a \in A$ given. This leads to a homomorphism from $A((T))$ into $B((T))$ as modules over $k$, as in Section 8.4. More precisely, one can verify that this is a homomorphism from $A((T))$ into $B((T))$ as modules over $k((T))$. One can also look at this in terms of $k$-linear mappings satisfying (8.5.1) and (8.5.3).

Similarly, let $C$ be a third module over $k$, and let $\beta$ be a mapping from $A \times B$ into $C((T))$ that is bilinear over $k$ and has poles of bounded order. One can check that the extension of $\beta$ to a mapping from $A((T)) \times B((T))$ into $C((T))$ defined in Section 8.6 is bilinear over $k((T))$. One can also look at this in terms of mappings from $A((T)) \times B((T))$ into $C((T))$ that are bilinear over $k$ and satisfy (8.7.1) and (8.7.3). In particular, if $A$ is an algebra over $k$ in the strict sense, then $A((T))$ may be considered as an algebra over $k((T))$ in the strict sense.

Let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. If $f(T), g(T) \in A((T))$, then (8.8.2) should be expressed as

$$h_n = \sum_{j=-\infty}^{\infty} [f_j, g_{n-j}]_A$$

for each $n \in Z$, so that (8.8.5) corresponds to

$$[f(T), g(T)]_{A((T))} = h(T) = \sum_{n=-\infty}^{\infty} h_n T^n.$$

As in Section 8.6, $[f(T), f(T)]_{A((T))} = 0$ for every $f(T) \in A((T))$, because $[a, a]_A = 0$ for every $a \in A$. One can also check that $[\cdot, \cdot]_{A((T))}$ satisfies the Jacobi identity on $A((T))$, using the Jacobi identity for $[\cdot, \cdot]_A$ on $A$. Thus $A((T))$ is a Lie algebra with respect to (8.8.9) over $k$, and in fact over $k((T))$.

8.9 Absolute values on $k((T))$

Let $k$ be a field, and let $T$ be an indeterminate. If $f(T)$ is a nonzero element of $k((T))$, then $f(T)$ is invertible in $k((T))$, because nonzero elements of $k$ are invertible in $k$, and using the remark about invertibility in the previous section. Thus $k((T))$ is a field. Let $f(T) = \sum_{j=-\infty}^{\infty} f_j T^j \in k((T))$ be given. If $f(T) \neq 0$, then there is a unique minimal integer $j_0(f(T))$ such that $f_{j_0(f(T))} \neq 0$, which is to say that $f_j = 0$ when $j < j_0(f(T))$. If $f(T) = 0$, then it is convenient to put $j_0(f) = +\infty$. Observe that

$$j_0(f(T) + g(T)) \geq \min(j_0(f(T)), j_0(g(T)))$$

and

$$j_0(f(T) g(T)) = j_0(f(T)) + j_0(g(T))$$
for every \( f(T), g(T) \in k((T)) \), with suitable interpretations when any of these terms is \( +\infty \).

Let \( r \) be a positive real number with \( r \leq 1 \). If \( f(T) \in k((T)) \), then put

\[
|f(T)|_r = r^{\text{ht}(f(T))}
\]

when \( f(T) \neq 0 \), and \( |f(T)|_r = 0 \) when \( f(T) = 0 \). Using (8.9.1) and (8.9.2), we get that

\[
|f(T) + g(T)|_r \leq \max(|f(T)|_r, |g(T)|_r)
\]

and

\[
|f(T) g(T)|_r = |f(T)|_r |g(T)|_r
\]

for every \( f(T), g(T) \in k((T)) \). Thus \( |f(T)|_r \) defines an ultrametric absolute value function on \( k((T)) \), which is the same as the trivial absolute value function on \( k((T)) \) when \( r = 1 \). If \( a \) is a positive real number, then \( 0 < r^a \leq 1 \), and

\[
|f(T)|_{r^a} = |f(T)|_r^a
\]

for every \( f(T) \in k((T)) \).

It follows that

\[
d_r(f(T), g(T)) = |f(T) - g(T)|_r
\]

is an ultrametric on \( k((T)) \), which is the discrete metric when \( r = 1 \). Let us suppose from now on in this section that \( r < 1 \). Let \( l \in \mathbb{Z} \), then the closed ball in \( k((T)) \) centered at 0 with radius \( r^l \) with respect to (8.9.7) is given by

\[
B(0, r^l) = \{ f(T) \in k((T)) : |f(T)|_r \leq r^l \} = \{ f(T) \in k((T)) : j_0(f(T)) \geq l \} = (k[[T]]) T^l.
\]

Let \( l \in \mathbb{Z} \) be given, so that (8.9.8) can be identified with the space of \( k \)-valued functions on the set of integers \( j \geq l \). This may be considered as the Cartesian product of copies of \( k \) indexed by integers \( j \geq l \). One can check that the topology determined on (8.9.8) by the restriction of the ultrametric (8.9.7) to (8.9.8) corresponds to the product topology on the Cartesian product just mentioned, using the discrete topology on \( k \).

One can verify that \( k((T)) \) is complete with respect to the ultrametric (8.9.7), as follows. Any Cauchy sequence in \( k((T)) \) with respect to (8.9.7) is contained in (8.9.8) for some \( l \in \mathbb{Z} \), because a Cauchy sequence in any metric space is bounded. It is easy to see that for each \( j \in \mathbb{Z} \), the corresponding sequence of coefficients of \( T^j \) in the terms of the Cauchy sequence is eventually constant, as a sequence of elements of \( k \). This leads to an element of (8.9.8), for which the coefficient of \( T^j \) is the eventual constant value of the sequence in \( k \) just mentioned, for each \( j \in \mathbb{Z} \). The given Cauchy sequence in \( k((T)) \) converges to this element of (8.9.8) with respect to the ultrametric (8.9.7), by the description of the topology determined on (8.9.8) by the restriction of the ultrametric (8.9.7) in the preceding paragraph.
8.10  Formal series and algebra homomorphisms

Let $k$ be a commutative ring with a multiplicative identity element, and let $T$ be an indeterminate. Also let $A$, $B$ be algebras over $k$ in the strict sense, where multiplication of $x$, $y$ is expressed as $xy$. Remember that multiplication on $A$ and $B$ can be extended to $A((T))$ and $B((T))$, respectively, so that $A((T))$ and $B((T))$ become algebras in the strict sense over $k((T))$, as in Section 8.8. In particular, they may be considered as algebras over $k$.

Let $\phi$ be a homomorphism from $A$ into $B((T))$, as modules over $k$ for the moment. As in Section 8.3, $\phi$ can be expressed as

\[
\phi(a) = \sum_{j=-\infty}^{\infty} \phi_j(a) T^j
\]

for each $a \in A$, where $\phi_j$ is a module homomorphism from $A$ into $B$ for every $j \in \mathbb{Z}$, and for every $a \in A$ we have that $\phi_j(a) = 0$ for all but finitely many $j < 0$. Of course, $\phi$ is a homomorphism from $A$ into $B((T))$ as algebras over $k$ when

\[
\phi(aa') = \phi(a) \phi(a')
\]

for every $a, a' \in A$. Let $a, a' \in A$ be given, so that

\[
\phi(aa') = \sum_{n=-\infty}^{\infty} \phi_n(aa') T^n,
\]

as in (8.10.1), where $\phi_n(aa') = 0$ for all but finitely many $n < 0$. We also have that

\[
\phi(a) \phi(a') = \left( \sum_{j=-\infty}^{\infty} \phi_j(a) T^j \right) \left( \sum_{l=-\infty}^{\infty} \phi_l(a') T^l \right) = \sum_{n=-\infty}^{\infty} \left( \sum_{j=-\infty}^{\infty} \phi_j(a) \phi_{n-j}(a') \right) T^n,
\]

as in Section 8.8. Remember that for each $n \in \mathbb{Z}$, $\phi_j(a) \phi_{n-j}(a') = 0$ for all but finitely many $j \in \mathbb{Z}$, and that for all but finitely many $n < 0$, this condition holds for every $j \in \mathbb{Z}$. Comparing (8.10.3) and (8.10.4), we get that (8.10.2) holds if and only if

\[
\phi_n(aa') = \sum_{j=-\infty}^{\infty} \phi_j(a) \phi_{n-j}(a')
\]

for every $n \in \mathbb{Z}$.

Suppose that $\phi$ has poles of bounded order on $A$, so that there is an integer $r(\phi)$ such that $\phi_j = 0$ on $A$ when $j < r(\phi)$. Let $f(T) = \sum_{j=j_0}^{\infty} f_j T^j$ be an element of $A((T))$. As in Section 8.4, $\phi$ can be extended to a module homomorphism from $A((T))$ into $B((T))$, by putting

\[
\phi(f(T)) = \sum_{j=j_0}^{\infty} \phi(f_j) T^j,
\]
where the sum on the right is defined as an element of $B((T))$ as in Section 8.2. Let $g(T) = \sum_{i=0}^{\infty} g_i T^i$ be another element of $A((T))$, so that

\[(8.10.7) \quad \phi(g(T)) = \sum_{l=0}^{\infty} \phi(g_l) T^l.\]

Remember that $f(T)g(T) = h(T) = \sum_{n=j_0+l_0}^{\infty} h_n T^n$ is defined in $A((T))$ by putting

\[(8.10.8) \quad h_n = \sum_{j=-\infty}^{\infty} f_j g_{n-j}\]

for each $n$, which reduces to a finite sum in $A$. Thus

\[(8.10.9) \quad \phi(f(T)g(T)) = \phi(h(T)) = \sum_{n=j_0+l_0}^{\infty} \phi(h_n) T^n,\]

where the sum on the right is defined as an element of $B((T))$ as in Section 8.2 again. If $\phi$ is an algebra homomorphism from $A$ into $B((T))$, then

\[(8.10.10) \quad \phi(h_n) = \sum_{j=-\infty}^{\infty} \phi(f_j g_{n-j}) = \sum_{j=-\infty}^{\infty} \phi(f_j) \phi(g_{n-j})\]

for each $n$, where the sums reduce to finite sums in $B((T))$. We would like to say that

\[(8.10.11) \quad \sum_{n=j_0+l_0}^{\infty} \left( \sum_{j=-\infty}^{\infty} \phi(f_j) \phi(g_{n-j}) \right) T^n = \left( \sum_{j=j_0}^{\infty} \phi(f_j) T^j \right) \left( \sum_{l=l_0}^{\infty} \phi(g_l) T^l \right),\]

as elements of $B((T))$. If $\phi$ maps $A$ into $B$, then this follows from the definition of the extension of multiplication on $B$ to $B((T))$, as in Section 8.8. Otherwise, the $\phi(f_j)$’s, $j \geq j_0$, and $\phi(g_l)$’s, $l \geq l_0$, are elements of $B((T))$ with poles of bounded order, and (8.10.11) can be obtained from (8.7.7). It follows that

\[(8.10.12) \quad \phi(f(T)g(T)) = \phi(f(T)) \phi(g(T)),\]

so that the extension of $\phi$ to $A((T))$ is an algebra homomorphism as well.

There are analogous statements for opposite algebra homomorphisms, as usual.

### 8.11 Involutions and formal series

Let $k$ be a commutative ring with a multiplicative identity element, and let $T$ be an indeterminate. Also let $A$, $B$ be modules over $k$, and let $\phi$ be a module homomorphism from $A$ into $B$. If $f(T) = \sum_{j=-\infty}^{\infty} f_j T^j \in A((T))$, then

\[(8.11.1) \quad \phi(f(T)) = \sum_{j=-\infty}^{\infty} \phi(f_j) T^j\]
defines an element of $B((T))$. This defines a homomorphism from $A((T))$ into $B((T))$, as modules over $k((T))$. This may be seen as a simple case of the situation discussed in Section 8.4, by identifying $\phi$ with an element of $(\text{Hom}_k(A, B))((T))$. In particular, we have that
\[ (8.11.2) \quad \phi(f(T) T^r) = \phi(f(T)) T^r \]
for every $f(T) \in A((T))$ and $r \in \mathbb{Z}$. Of course, if $f(T)$ corresponds to an element of $A[T]$ or $A[[T]]$, then $\phi(f(T))$ corresponds to an element of $B[T]$ or $B[[T]]$, as appropriate. If $\phi$ is a one-to-one mapping from $A$ onto $B$, then (8.11.1) defines a one-to-one mapping from $A((T))$ onto $B((T))$.

Now let $A$, $B$ be algebras over $k$ in the strict sense, where multiplication of $x$, $y$ is expressed as $xy$. If $\phi$ is an algebra homomorphism from $A$ into $B$, then (8.11.1) defines an algebra homomorphism from $A((T))$ into $B((T))$, as in the previous section. Similarly, if $\phi$ is an opposite algebra homomorphism from $A$ into $B$, then (8.11.1) defines an opposite algebra homomorphism from $A((T))$ into $B((T))$.

Let $x \mapsto x^*$ be an algebra involution on $A$, and put
\[ (8.11.3) \quad f(T)^* = \sum_{j > 0} f_j^* T^{-j} \]
for every $f(T) \in A((T))$. This defines an algebra involution on $A((T))$, as before. Clearly $f(T) \in A((T))$ is self-adjoint with respect to this involution if and only if $f_j$ is self-adjoint in $A$ for every $j$. Similarly, $f(T)$ is anti-self-adjoint with respect to this involution if and only if $f_j$ is anti-self-adjoint in $A$ for every $j$.

If $k$ is the field $\mathbb{C}$ of complex numbers, then there are analogous statements for conjugate-linear mappings. More precisely, if $A$ and $B$ are vector spaces over $\mathbb{C}$, then $A((T))$ and $B((T))$ may be considered as vector spaces over $\mathbb{C}$ too. If $\phi$ is a conjugate-linear mapping from $A$ into $B$, then (8.11.1) defines a conjugate-linear mapping from $A((T))$ into $B((T))$, as vector spaces over $\mathbb{C}$.

If $k$ is the field $\mathbb{C}$ of complex numbers, then there are analogous statements for conjugate-linear mappings. More precisely, if $A$ and $B$ are vector spaces over $\mathbb{C}$, then $A((T))$ and $B((T))$ may be considered as vector spaces over $\mathbb{C}$ too. If $\phi$ is a conjugate-linear mapping from $A$ into $B$, then (8.11.1) defines a conjugate-linear mapping from $A((T))$ into $B((T))$, as conjugate-linear mappings are considered as real vector spaces, and that conjugate-linear mappings between complex vector spaces may be considered as real-linear mappings between the corresponding real vector spaces. This can be used to reduce statements about conjugate-linear mappings to the analogous statements for real-linear mappings, as before.

In particular, if $a(T) = \sum_{j > 0} a_j T^j \in \mathbb{C}((T))$, then
\[ (8.11.4) \quad \overline{a(T)} = \sum_{j > 0} \overline{a_j} T^{-j} \]
defines an element of $\mathbb{C}((T))$ too, and this defines a conjugate-linear automorphism of $\mathbb{C}((T))$ as an algebra over $\mathbb{C}$. If $a(T)$ corresponds to an element of $\mathbb{C}[T]$ or $\mathbb{C}[[T]]$, then $\overline{a(T)}$ corresponds to an element of $\mathbb{C}[T]$ or $\mathbb{C}[[T]]$, as appropriate. One can use this to define conjugate-linearity over $\mathbb{C}[T]$, $\mathbb{C}[[T]]$, and $\mathbb{C}((T))$. These conjugate-linearity conditions amount to ordinary conjugate-linearity over $\mathbb{C}$, together with the appropriate “real” linearity condition over $\mathbb{R}[T]$, $\mathbb{R}[[T]]$, or $\mathbb{R}((T))$, respectively.
8.12 Ordered rings

Let \( R \) be a ring with a nonzero multiplicative identity element \( e \). Of course, \( R \) may be considered as an associative algebra over \( \mathbb{Z} \), and in particular as a module over \( \mathbb{Z} \). Suppose that certain nonzero elements \( x \) of \( R \) have been designated as positive, which may be expressed by

\[
(8.12.1) \quad x > 0.
\]

We say that \( R \) is an ordered ring if the following conditions are satisfied. First, if \( x \) and \( y \) are positive elements of \( R \), then

\[
(8.12.2) \quad x + y > 0, \quad xy > 0.
\]

Second, if \( x \) is any nonzero element of \( R \), then either \( x \) or \( -x \) is positive. This corresponds to the definition on p261 of [18].

Let \( R \) be an ordered ring. If \( x, y \) are nonzero elements of \( R \), then \( xy \neq 0 \). More precisely,

\[
(8.12.3) \quad xy = (-x)(-y) > 0
\]

when \( x, y > 0 \), and when \( -x, -y > 0 \). Similarly,

\[
(8.12.4) \quad -xy = (-x)y = x(-y) > 0
\]

when \( -x, y > 0 \), and when \( x, -y > 0 \).

If \( R \) is commutative, then it follows that \( R \) is an integral domain. This corresponds to the definition of an ordered integral domain on p9 of [3]. Similarly, an ordered field is an ordered ring that is also a field, as in [3, 18].

Let \( R \) be an ordered ring again. Note that \( x \in R \) and \( -x \) cannot both be positive, because that would imply that \( x + (-x) = 0 \) is positive. If \( x \in R \) and \( x \neq 0 \), then

\[
(8.12.5) \quad x^2 > 0,
\]

because \( x \) or \( -x \) is positive, and \( x^2 = (-x)^2 \). Of course, this corresponds to (8.12.3), with \( y = x \). In particular,

\[
(8.12.6) \quad e = e^2 > 0
\]

in \( R \). If \( x, y \in R \), then put

\[
(8.12.7) \quad x < y
\]

when \( y - x > 0 \). This defines a linear ordering on \( R \), which is invariant under translations on \( R \).

Alternatively, one might start with a translation-invariant linear ordering on \( R \), and define \( x \in R \) to be positive when \( x > 0 \) with respect to this ordering. One can check that the sum of positive elements of \( R \) is positive in this situation. If products of positive elements of \( R \) are positive too, then \( R \) is an ordered ring. This is how ordered fields are defined on p7 in [19].

Clearly \( \mathbb{Z} \) is an ordered ring with respect to the standard ordering. In fact, this is the only ordering on \( \mathbb{Z} \) for which \( \mathbb{Z} \) is an ordered ring. More precisely,
if $\mathbb{Z}$ is an ordered ring with respect to some ordering, then 1 has to be positive in $\mathbb{Z}$ with respect to that ordering, by (8.12.6). This implies that all sums of 1 have to be positive with respect to this ordering on $\mathbb{Z}$. One can check that these are the only elements of $\mathbb{Z}$ that can be positive, so that this ordering on $\mathbb{Z}$ is the same as the usual one.

Let $R$ be an ordered ring, and let $R_0$ be a subring of $R$ that contains $e$. It is easy to see that $R_0$ is an ordered ring too, with respect to the restriction of the restriction of the ordering on $R$ to $R_0$.

### 8.13 Some additional features

Let $R$ be a ring, with multiplicative identity element $e$. If $n$ is a positive integer, then $n \cdot e$ is the sum of $n$ $e$'s in $R$, as usual. We can extend this to integers $n \leq 0$ in the obvious way, by putting $0 \cdot e = 0$ in $R$, and $n \cdot e = -((-n) \cdot e)$ when $n < 0$. This defines a ring homomorphism from $\mathbb{Z}$ into $R$.

Suppose that $R$ is an ordered ring. If $n \in \mathbb{Z}_+$, then

$$(8.13.1) \quad n \cdot e > 0$$

in $R$, by (8.12.6). Of course, this implies that

$$(8.13.2) \quad -(n \cdot e) = (-n) \cdot e > 0$$

when $-n \in \mathbb{Z}_+$. Thus $n \mapsto n \cdot e$ is an injective order-preserving mapping from $\mathbb{Z}$ into $R$, with respect to the standard ordering on $\mathbb{Z}$.

If $x \in R$, $x > 0$, and $x$ has a multiplicative inverse $x^{-1}$ in $R$, then

$$(8.13.3) \quad x^{-1} > 0.$$ 

Otherwise, if $-x^{-1} > 0$, then $-e = x(-x^{-1}) > 0$, contradicting (8.12.6).

If $x \in R$, then the absolute value $|x|$ of $x$ may be defined as an element of $R$ by

$$(8.13.4) \quad |x| = x \quad \text{when } x \geq 0$$

$$|x| = -x \quad \text{when } -x \geq 0,$$

as on p10 of [3], and p264 of [18]. Note that $|x| \geq 0$ and

$$(8.13.5) \quad -|x| \leq x \leq |x|$$

for every $x \in R$. One can check that

$$(8.13.6) \quad |xy| = |x||y|$$

and

$$(8.13.7) \quad |x + y| \leq |x| + |y|$$

for every $x, y \in R$. More precisely, (8.13.6) is basically the same as (8.12.3) and (8.12.4). To get (8.13.7), and one can use (8.13.5) and its analogue for $y$. 

Let \( T \) be an indeterminate, and let \( R((T)) \) be as before. More precisely, \( R \) may be considered as an associative algebra over \( \mathbb{Z} \), so that \( R((T)) \) is an associative algebra over \( \mathbb{Z} \) too. If \( f(T) \in R((T)) \) can be expressed as \( \sum_{j=j_0}^\infty f_j T^j \), where \( f_{j_0} > 0 \) in \( R \), then let us say that \( f(T) \) is positive as an element of \( R((T)) \).

Let us check that this makes \( R((T)) \) into an ordered ring, as in the discussion on p284-5 in [18]. Note that the elements of \( R((T)) \) are called extended formal power series in [18].

Let \( f(T) \) be as in the preceding paragraph, and let \( g(T) = \sum_{l=0}^\infty g_l T^l \) be another positive element of \( R((T)) \), with \( g_{l_0} > 0 \). Put \( h(T) = f(T) g(T) \), so that \( h(T) = \sum_{n=j_0+l_0}^\infty h_n T^n \), with

\[
(8.13.8) \quad h_{j_0+l_0} = f_{j_0} g_{l_0}.
\]

This implies that \( h_{j_0+l_0} > 0 \) in \( R \), so that \( h(T) > 0 \) in \( R((T)) \).

One can verify that

\[
(8.13.9) \quad f(T) + g(T) > 0
\]

in \( R((T)) \), directly from the definitions. More precisely, if \( l_0 > j_0 \), then (8.13.9) holds when \( f_{j_0} > 0 \) in \( R \), without additional conditions on \( g(T) \). Similarly, if \( j_0 > l_0 \), then (8.13.9) holds when \( g_{l_0} > 0 \) in \( R \), without additional conditions on \( f(T) \).

If \( a(T) \) is any nonzero element of \( R((T)) \), then \( a(T) \) can be expressed as \( \sum_{r=r_0}^\infty a_r T^r \), where \( a_{r_0} \neq 0 \). Because \( R \) is an ordered ring, either \( a_{r_0} > 0 \) or \( -a_{r_0} > 0 \) in \( R \). This implies that \( a(T) > 0 \) or \( -a(T) > 0 \) in \( R((T)) \), as appropriate.

### 8.14 Ordered fields

Note that \( \mathbb{Q} \) is an ordered field with respect to the standard ordering. One can check that this is the only ordering on \( \mathbb{Q} \) for which \( \mathbb{Q} \) is an ordered field. Indeed, if \( \mathbb{Q} \) is an ordered field with respect to some ordering, then every \( n \in \mathbb{Z}_+ \) is positive with respect to this ordering on \( \mathbb{Q} \), as before. This implies that \( 1/n \) is positive with respect to this ordering on \( \mathbb{Q} \), as in the previous section, and hence that quotients of elements of \( \mathbb{Z}_+ \) are positive with respect to this ordering on \( \mathbb{Q} \). One can verify that these are the only elements of \( \mathbb{Q} \) that can be positive, so that this ordering on \( \mathbb{Q} \) is the standard ordering.

Let \( k \) be an ordered field. Note that \( k \) has characteristic 0, because sums of 1 are positive in \( k \). The usual homomorphism from \( \mathbb{Z} \) into \( k \) extends to a field isomorphism from \( \mathbb{Q} \) onto a subfield of \( k \). This isomorphism is also compatible with the standard ordering on \( \mathbb{Q} \), as in the preceding paragraph. This corresponds to the corollary on p266 of [18].

The classical version of the archimedean property can be stated for \( k \) as follows: if \( x, y \) are positive elements of \( k \), then there is a positive integer \( n \) such that

\[
(8.14.1) \quad n \cdot x > y.
\]
Of course, \( \mathbb{R} \) has the archimedean property with respect to the standard ordering. If \( k \) has the archimedean property, then every subfield of \( k \) has the archimedean property.

Let \( k_1 \) and \( k_2 \) be ordered fields. To say that \( k_1 \) and \( k_2 \) are isomorphic as ordered fields means that there is a field isomorphism from \( k_1 \) onto \( k_2 \) that preserves order as well. Clearly the archimedean property is invariant under order-preserving field isomorphisms.

Suppose that \( k_1 \) is an ordered field with the archimedean property. It is well known that \( k_1 \) is isomorphic as an ordered field to a subfield of \( \mathbb{R} \), with the ordering induced by the standard ordering on \( \mathbb{R} \). This corresponds to Exercise 10 on p286 of [18].

Let \( k \) be an ordered field again, and let \( T \) be an indeterminate. Thus \( k((T)) \) is an ordered field with respect to the ordering obtained from the one on \( k \) described in the previous section. It is easy to see that \( k((T)) \) does not have the archimedean property with respect to this ordering, even if \( k \) has the archimedean property, as on p285 of [18]. More precisely, \( T \) is a positive element of \( k((T)) \), because \( 1 > 0 \) in \( k \). However,

\[
(8.14.2) \quad n \cdot T < 1
\]

for every positive integer \( n \).

Some topics related to inner products on vector spaces over ordered fields will be discussed in Section 11.9.
Part II

Solvability, nilpotence, and semisimplicity
Chapter 9

Solvability and nilpotence

9.1 Some basic isomorphism theorems

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an algebra over $k$ in the strict sense. If $A_0$ is a two-sided ideal in $A$, then the quotient $A/A_0$ can be defined as an algebra over $k$ in the strict sense too, as in Section 2.11. Let $q_0$ be the canonical quotient mapping from $A$ onto $A/A_0$, which is an algebra homomorphism.

Suppose that $B$ is another algebra over $k$ in the strict sense, and that $\phi$ is an algebra homomorphism from $A$ into $B$. The kernel $\ker \phi$ of $\phi$ is a two-sided ideal in $A$, as mentioned in Section 2.11. If $A_0 \subseteq \ker \phi$, then there is a unique algebra homomorphism $\psi$ from $A/A_0$ onto $\phi(A) \subseteq B$ such that

$$\psi \circ q_0 = \phi.$$  \hspace{1cm} (9.1.1)

If $A_0 = \ker \phi$, then $\psi$ is injective.

Let $A_1, A_2$ be two-sided ideals in $A$ such that $A_1 \subseteq A_2$, and let $q_1, q_2$ be the canonical quotient mappings from $A$ onto $A/A_1, A/A_2$, respectively. There is a natural algebra homomorphism $\Psi$ from $A/A_1$ onto $A/A_2$ such that

$$\Psi \circ q_1 = q_2.$$  \hspace{1cm} (9.1.2)

as in the preceding paragraph.

Note that $A_1$ may be considered as a two-sided ideal in $A_2$, so that $A_2/A_1$ can be defined as an algebra over $k$ in the strict sense. Of course, $A_2/A_1$ is essentially the same as $q_1(A_2) \subseteq A/A_1$. It is easy to see that $q_1(A_2)$ is a two-sided ideal in $A/A_1$, because $A_2$ is a two-sided ideal in $A$.

The kernel of $\Psi$ is equal to $q_1(A_2)$, by construction. Thus $\Psi$ can be identified with the quotient mapping from $A/A_1$ onto $(A/A_1)/q_1(A_2)$, which is the same as $(A/A_1)/(A_2/A_1)$. This leads to a natural algebra isomorphism between this quotient and $A/A_2$.

Let $A_3, A_4$ be two-sided ideals in $A$, and observe that $A_3 \cap A_4$ is a two-sided ideal in $A$ as well. Remember that $A_3 + A_4$ is a two-sided ideal in $A$ too, as in
Section 7.1. Let \( q \) be the canonical quotient mapping from \( A \) onto \( A/A_4 \). The restriction of \( q \) to \( A_3 + A_4 \) is essentially the same as the canonical quotient mapping from \( A_3 + A_4 \) onto \( (A_3 + A_4)/A_4 \), as before.

Observe that \( q(A_3) = q(A_3 + A_4) \). The kernel of the restriction of \( q \) to \( A_3 \) is equal to \( A_3 \cap A_4 \). The restriction of \( q \) to \( A_3 \), as an algebra homomorphism from \( A_3 \) onto \( q(A_3) \), can be identified with the quotient mapping from \( A_3 \) onto \( A_3/(A_3 \cap A_4) \). This leads to a natural algebra isomorphism between this quotient and \( q(A_3 + A_4) \), which is essentially the same as \( (A_3 + A_4)/A_4 \).

These isomorphism theorems are stated for Lie algebras (over fields) on p7-8 of [13].

9.2 Products of ideals

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( A \) be an algebra over \( k \) in the strict sense, where multiplication of \( a, b \in A \) is expressed as \( a \cdot b \). Also let \( A_1 \) and \( A_2 \) be submodules of \( A \) as a module over \( k \). The product \( A_1 \cdot A_2 \) of \( A_1 \) and \( A_2 \) is defined to be the subset of \( A \) consisting of all finite sums of elements of \( A \) of the form \( a_1 \cdot a_2 \), where \( a_1 \in A_1 \) and \( a_2 \in A_2 \). It is easy to see that \( A_1 \cdot A_2 \) is a submodule of \( A \) as well, as a module over \( k \). If multiplication on \( A \) is commutative or anti-commutative, then

\[
(9.2.1) \quad A_1 \cdot A_2 = A_2 \cdot A_1.
\]

If \( A_1 \) is a right ideal in \( A \), then

\[
(9.2.2) \quad A_1 \cdot A_2 \subseteq A_1.
\]

Similarly, if \( A_2 \) is a left ideal in \( A \), then

\[
(9.2.3) \quad A_1 \cdot A_2 \subseteq A_2.
\]

Suppose for the moment that \( A \) is an associative algebra over \( k \). If \( A_1 \) is a left ideal in \( A \), then \( A_1 \cdot A_2 \) is a left ideal in \( A \) too. If \( A_2 \) is a right ideal in \( A \), then \( A_1 \cdot A_2 \) is a right ideal in \( A \).

Let \( B \) be a subalgebra of \( A \), as an algebra over \( k \) in the strict sense. If \( B_1 \) and \( B_2 \) are submodules of \( B \), as a module over \( k \), then \( B_1 \) and \( B_2 \) are submodules of \( A \) too, so that \( [B_1, B_2] \) can be defined as a submodule of \( A \) as before. This is the same as defining \( [B_1, B_2] \) as a submodule of \( B \) in the analogous way.

Let \( C \) be another algebra over \( k \) in the strict sense, and let \( \phi \) be an algebra homomorphism from \( A \) into \( C \). If \( A_1 \) and \( A_2 \) are submodules of \( A \), as a module over \( k \), then \( \phi(A_1) \) and \( \phi(A_2) \) are submodules of \( C \), so that \( \phi(A_1) \cdot \phi(A_2) \) can be defined as a submodule of \( C \) as before. Observe that

\[
(9.2.4) \quad \phi(A_1 \cdot A_2) = \phi(A_1) \cdot \phi(A_2).
\]

Suppose now that \( (A, [\cdot, \cdot]_A) \) is a Lie algebra over \( k \). In this situation, \( A_1 \cdot A_2 \) may be denoted \([A_1, A_2]_A\), and consists of finite sums of elements of \( A \) of the
form \([a_1, a_2]_A\), where \(a_1 \in A_1\) and \(a_2 \in A_2\), as before. Note that

\[
[A_1, A_2] = [A_2, A_1],
\]

as in (9.2.1), because of anticommutativity of the Lie bracket on \(A\). If \(A_1\) and \(A_2\) are ideals in \(A\), then it is easy to see that \([A_1, A_2]\) is an ideal in \(A\), using the Jacobi identity.

In particular, we can apply this to \(A_1 = A_2 = A\), to get that \([A;A]\) is an ideal in \(A\). This is known as the derived algebra associated to \(A\). By construction, \(A=[A;A]\) is commutative as a Lie algebra over \(k\).

Let \(A_0\) be an ideal in \(A\), and suppose that \(A=A_0\) is commutative as a Lie algebra over \(k\). If \(a_1, a_2 \in A\), then the image of \([a_1, a_2]_A\) in \(A=A_0\) is the same as the Lie bracket of the images of \(a_1\) and \(a_2\) in \(A/A_0\), which is equal to 0. This means that \([a_1, a_2] \in A_0\), so that

\[
[A;A] \subseteq A_0.
\]

Equivalently, if \(\phi\) is a homomorphism from \(A\) into a commutative Lie algebra over \(k\), then one can take \(A_0\) to be the kernel of \(\phi\).

Let \(B\) be a Lie subalgebra of \(A\), so that \(B\) may be considered as a Lie algebra over \(k\) too. If \(B_1, B_2\) are ideals in \(B\), then \([B_1, B_2]\) is an ideal in \(B\), as before. In particular, \([B;B]\) is an ideal in \(B\), and

\[
[B;B] \subseteq [A;A].
\]

### 9.3 Solvable Lie algebras

Let \(k\) be a commutative ring with a multiplicative identity element, and let \((A, [\cdot, \cdot])_A\) be a Lie algebra over \(k\). If \(j\) is a nonnegative integer, then \(A^{(j)}\) is defined inductively by putting \(A^{(0)} = A\), \(A^{(1)} = [A;A]\), and

\[
A^{(j+1)} = [A^{(j)}, A^{(j)}]
\]

for each \(j \geq 0\), where the right side is as defined in the preceding section. It is easy to see that \(A^{(j)}\) is an ideal in \(A\) for every \(j \geq 0\), using induction and a remark in the previous section. This sequence of ideals is called the derived series of \(A\). Here we use the notation on p10 of [13], while on p35 of [24] the notation \(D^j A\) is used for \(A^{(j-1)}\) when \(j \geq 1\).

Observe that

\[
A^{(j+1)} \subseteq A^{(j)}
\]

for every \(j \geq 0\). Of course, \(A^{(1)} = \{0\}\) exactly when \(A\) is commutative as a Lie algebra. If \(A^{(j)} = \{0\}\) for some \(j \geq 0\), then \(A\) is said to be solvable as a Lie algebra.

If \(B\) is a Lie subalgebra of \(A\), then \(B^{(j)}\) can be defined as an ideal in \(B\) for each \(j \geq 0\) as before. One can check that

\[
B^{(j)} \subseteq A^{(j)}
\]
for every $j \geq 0$, using (9.2.7) and induction. In particular, if $A$ is solvable, then $B$ is solvable.

By construction,
\[ A^{(j+l)} = (A^{(j)})^{(l)} \]
for every $j, l \geq 0$. If $A^{(j)}$ is solvable for some $j \geq 0$, then it follows that $A$ is solvable.

Let $C$ be another Lie algebra over $k$, and let $\phi$ be a Lie algebra homomorphism from $A$ onto $C$. One can verify that
\[ \phi(A^{(j)}) = C^{(j)} \]
for every $j \geq 0$, using (9.2.4) and induction. If $A$ is solvable, then it follows that $C$ is solvable.

Suppose that $C$ is solvable, so that $C^{(n)} = \{0\}$ for some nonnegative integer $n$. This implies that
\[ \phi(A^{(n)}) = C^{(n)} = \{0\}, \]
as in (9.3.5), which means that
\[ A^{(n)} \subseteq \ker \phi. \]

Suppose that the kernel of $\phi$ is solvable as a Lie algebra over $k$ as well, so that $(\ker \phi)^{(l)} = \{0\}$ for some nonnegative integer $l$. Under these conditions, we get that
\[ A^{(n+l)} = (A^{(n)})^{(l)} \subseteq (\ker \phi)^{(l)} = \{0\}, \]
and hence that $A$ is solvable.

These properties correspond to parts (a) and (b) of the proposition on p11 of [13], and to part of Exercise 1 on p43 of [24]. Alternatively, let $A_0, A_1, \ldots, A_n$ be finitely many Lie subalgebras of $A$, with $A_0 = A$, and $A_{j+1}$ an ideal in $A_j$ for $j = 0, \ldots, n-1$. Suppose that $A_j/A_{j+1}$ is commutative as a Lie algebra for each $j = 0, \ldots, n-1$, which is the same as saying that
\[ [A_j, A_j] \subseteq A_{j+1} \]
for every $j = 0, \ldots, n-1$. This implies that
\[ A^{(j)} \subseteq A_j \]
for each $j = 0, 1, \ldots, n$, by induction. In particular, if $A_n = \{0\}$, then it follows that $A$ is solvable. Conversely, if $A$ is solvable, then $A^{(n)} = \{0\}$ for some $n \geq 0$, and one can simply take $A_j = A^{(j)}$ for $j = 0, 1, \ldots, n$. This characterization of solvability is mentioned on p35-6 of [24], and in Exercise 2 on p14 of [13].

**9.4 The solvable radical**

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot])$ be a Lie algebra over $k$. Also let $A_1$ and $A_2$ be ideals in $A$, so that
9.4. THE SOLVABLE RADICAL

\( A_1 \cap A_2 \) and \( A_1 + A_2 \) are ideals in \( A \) as well. If \( A_1 \) and \( A_2 \) are solvable as Lie algebras over \( k \), then it is well known that \( A_1 + A_2 \) is solvable as a Lie algebra too. To see this, it suffices to check that \( (A_1 + A_2)/A_2 \) is solvable as a Lie algebra over \( k \), because \( A_2 \) is solvable, as in the previous section. Remember that \( (A_1 + A_2)/A_2 \) is isomorphic to \( A_1/(A_1 \cap A_2) \), as in Section 9.1. Of course, \( A_1/(A_1 \cap A_2) \) is solvable as a Lie algebra, because \( A_1 \) is solvable, as in the previous section. Thus \( (A_1 + A_2)/A_2 \) is solvable, as desired. This is part (c) of the proposition on p11 of [13], which is mentioned on p44 of [24].

One often considers Lie algebras \( A \) over a field, with finite dimension as a vector space over the field. In this case, it is easy to see that there is a maximal solvable ideal in \( A \), by taking a solvable ideal in \( A \) of maximal dimension, as a linear subspace of \( A \). The remarks in the previous paragraph imply that a maximal solvable ideal in \( A \) is unique, and in fact contains every other solvable ideal in \( A \). This maximal solvable ideal is called the (solvable) radical of \( A \), and may be denoted \( \text{Rad} A \).

Let \( A \) be a Lie algebra over a commutative ring \( k \) with a multiplicative identity element again. It may still happen that \( A \) has a maximal solvable ideal, which can still be called the radical of \( A \), and denoted \( \text{Rad} A \). In particular, \( A \) has a maximal solvable ideal when solvable ideals in \( A \) satisfy an ascending chain condition. If a maximal solvable ideal exists, then it is unique, and contains all other solvable ideals in \( A \), as before. If

\[
(9.4.1) \quad \text{Rad} A = \{0\},
\]

then \( A \) may be called semisimple as a Lie algebra. More precisely, let us say that \( A \) is semisimple as a Lie algebra if \( \{0\} \) is the only solvable ideal in \( A \), in which case it is automatically maximal. Equivalently, \( A \) is not semisimple when \( A \) contains a nonzero solvable ideal, without asking for a maximal solvable ideal.

Suppose that \( A \) is not semisimple, and let \( B \) be a nonzero solvable ideal in \( A \). Note that the derived subalgebra \( B^{(1)} \) is an ideal in \( A \) for every nonnegative integer \( j \). This follows from a remark in Section 9.2 when \( j = 1 \), and can be verified using induction otherwise. Because \( B \neq \{0\} \) is solvable, there is a nonnegative integer \( j_0 \) such that \( B^{(j_0)} \neq \{0\} \) and \( B^{(j_0+1)} = \{0\} \). This means that \( B^{(j_0)} \) is commutative as a Lie algebra, since \( [B^{(j_0)}, B^{(j_0)}] = B^{(j_0+1)} = \{0\} \). Of course, if \( A \) has a nonzero ideal that is commutative as a Lie algebra, then \( A \) is not semisimple, because commutative Lie algebras are solvable. Thus \( A \) is not semisimple exactly when \( A \) has a nonzero ideal that is commutative as a Lie algebra, as on p22 of [13] and p44 of [24].

Suppose that the radical of \( A \) exists, so that the quotient \( A/\text{Rad} A \) can be defined as a Lie algebra over \( k \), and let \( q \) be the canonical quotient mapping from \( A \) onto \( A/\text{Rad} A \). In this situation, \( A/\text{Rad} A \) is automatically semisimple, as on p11 of [13]. Indeed, if \( C \) is any ideal in \( A/\text{Rad} A \), then \( q^{-1}(C) \) is an ideal in \( A \). If \( C \) is solvable as a Lie algebra over \( k \), then \( q^{-1}(C) \) is solvable as a Lie algebra too, because \( \text{Rad} A \) is solvable, as in the previous section. This implies that \( q^{-1}(C) = \text{Rad} A \), so that \( C = \{0\} \), as desired.
9.5 Nilpotent Lie algebras

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A,[,]_A)$ be a Lie algebra over $k$. If $j$ is a nonnegative integer, then $A^j$ is defined inductively by putting $A^0 = A$, $A^1 = [A,A]$, and
\begin{equation}
A^{j+1} = [A,A^j]
\end{equation}
for every $j \geq 0$, where the right side is as defined in Section 9.2. One can check that $A^j$ is an ideal in $A$ for every $j \geq 0$, using induction and a remark in Section 9.2. This sequence of ideals is called the descending central series or lower central series in $A$. This uses the notation on p11 in [13], and the notation $C^jA$ is used on p32 of [24] for $A^{j-1}$ when $j \geq 1$.

The fact that $A^j$ is an ideal in $A$ says exactly that
\begin{equation}
A^{j+1} \subseteq A^j
\end{equation}
for every $j \geq 0$. If $A^j = \{0\}$ for some $j \geq 0$, then $A$ is said to be nilpotent as a Lie algebra.

Observe that
\begin{equation}
A^{(j)} \subseteq A^j
\end{equation}
for every $j \geq 0$, with equality when $j = 0,1$. It follows that nilpotent Lie algebras are solvable. If $A$ is a commutative Lie algebra, then $A^1 = A^{(1)} = \{0\}$, and hence $A$ is nilpotent.

If $B$ is a Lie subalgebra of $A$, then $B^j$ can be defined as an ideal in $B$ in the same way as before, so that $B^0 = B$ and
\begin{equation}
B^{j+1} = [B,B^j]
\end{equation}
for every $j \geq 0$. It is easy to see that
\begin{equation}
B^j \subseteq A^j
\end{equation}
for every $j \geq 0$, by induction. If $A$ is nilpotent, then it follows that $B$ is nilpotent.

Let $\phi$ be a Lie algebra homomorphism from $A$ onto another Lie algebra $C$ over $k$. One can check that
\begin{equation}
\phi(A^j) = C^j
\end{equation}
for every $j \geq 0$, using (9.2.4) and induction. If $A$ is nilpotent, then it follows that $C$ is nilpotent as well.

Remember that $Z(A)$ is the center of $A$ as a Lie algebra, as in Section 7.6. If $B$ is a submodule of $A$, as a module over $k$, then
\begin{equation}
B \subseteq Z(A)
\end{equation}
if and only if
\begin{equation}
[A,B] = \{0\}.
\end{equation}
If $A \subseteq Z(A)$ for some $j \geq 0$, then it follows that

\[(9.5.9) \quad A^{j+1} = [A, A^j] = \{0\},\]

so that $A$ is nilpotent as a Lie algebra. If $A/Z(A)$ is nilpotent as a Lie algebra, then $(9.5.9)$ holds for some $j \geq 0$, because of $(9.5.6)$. This implies that $A$ is nilpotent as a Lie algebra, as before.

Let $B$ be a submodule of $A$ again, as a module over $k$, and let $B_0$ be an ideal in $A$. It is easy to see that

\[(9.5.11) \quad [A, B] \subseteq B_0\]

if and only if the image of $B$ in $A/B_0$ is contained in the center of $A/B_0$, as a Lie algebra.

Let $A_0, A_1, \ldots, A_n$ be finitely many ideals in $A$, with $A_0 = A$, and

\[(9.5.12) \quad A_{j+1} \subseteq A_j\]

for each $j = 0, \ldots, n - 1$. Suppose that $A_j/A_{j+1}$ is contained in the center of $A/A_{j+1}$ for each $j = 0, \ldots, n - 1$, which is the same as saying that

\[(9.5.13) \quad [A, A_j] \subseteq A_{j+1}\]

for every $j = 0, \ldots, n-1$, as in the preceding paragraph. Under these conditions, we get that

\[(9.5.14) \quad A^j \subseteq A_j\]

for each $j = 0, \ldots, n$, by induction. If $A_n = \{0\}$, then it follows that $A$ is nilpotent as a Lie algebra. Conversely, if $A$ is nilpotent as a Lie algebra, then $A^n = \{0\}$ for some nonnegative integer $n$, and one can take $A_j = A^j$ for $j = 0, \ldots, n$.

These basic properties of nilpotent Lie algebras correspond to parts (a) and (b) of the proposition on p12 of [13], part of Theorem 2.1 on p32 in [24], and part of Exercise 1 on p43 of [24].

Suppose that $A$ is nilpotent as a Lie algebra, and that $A \neq \{0\}$. Let $j$ be the largest nonnegative integer such that $A^j \neq \{0\}$, which means that $A^{j+1} = \{0\}$. This implies that $A^j \subseteq Z(A)$, and in particular that $Z(A) \neq \{0\}$. This is part (c) of the proposition on p12 of [13].

### 9.6 Two-dimensional Lie algebras

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot], A)$ be a Lie algebra over $k$. If $A$ is generated, as a module over $k$, by a single element, then it is easy to see that $A$ is commutative as a Lie algebra. Suppose now that $A$ is generated by $a_0, b_0 \in A$, as a module over $k$, so that every element of $A$ can be expressed as

\[(9.6.1) \quad \alpha a_0 + \beta b_0\]
for some $\alpha, \beta \in k$. This implies that $A^{(1)} = [A, A]$ consists of elements of $A$ of the form

\begin{equation}
\gamma [a_0, b_0]_A,
\end{equation}

where $\gamma \in k$. It follows that $A^{(2)} = [A^{(1)}, A^{(1)}] = \{0\}$, and in particular that $A$ is solvable.

By hypothesis,

\begin{equation}
[a_0, b_0]_A = \alpha_2 a_0 + \beta_0 b_0
\end{equation}

for some $\alpha_0, \beta_0 \in k$. Of course, if $[a_0, b_0]_A = 0$, then $A$ is commutative as a Lie algebra. Suppose that $[a_0, b_0]_A \neq 0$, and that (9.6.1) is not equal to 0 in $A$ when $\alpha, \beta \in k$ satisfy $\alpha \neq 0$ or $\beta \neq 0$. If $k$ has no nonzero nilpotent elements, then one can check that $A$ is not nilpotent as a Lie algebra. This corresponds to the first part of Exercise 5 on p14 of [13], and part of Exercise 2 on p43 of [24]. Suppose for the moment that $k$ is a field, so that $A$ is a two-dimensional vector space over $k$. One can choose a basis $a, b$ for $A$ such that

\begin{equation}
[a, b]_A = a,
\end{equation}

as on p5 of [13], and the other part of Exercise 2 on p43 of [24].

Let $A$ be a module over $k$, and let $[a, b]_A$ be a mapping from $A \times A$ into $A$ that is bilinear over $k$ and satisfies

\begin{equation}
[a, a]_A = 0
\end{equation}

for every $a \in A$. In order for $[\cdot, \cdot]_A$ to define a Lie bracket on $A$, one should verify that the Jacobi identity holds for any triple of elements $x, y, \text{ and } z$ of $A$. If $x = y = z$, then each of the three terms in the Jacobi identity is equal to 0, because of (9.6.5). If any two of $x, y, \text{ and } z$ are the same element of $A$, then one of the terms in the Jacobi identity is automatically equal to 0, by (9.6.5) again. In this case, the Jacobi identity can be obtained using the antisymmetry of $[\cdot, \cdot]_A$, which follows from (9.6.5), as usual.

Suppose that $A$ is generated as a module over $k$ by $a_0, b_0 \in A$, so that every element of $A$ can be expressed as in (9.6.1). In order to show that $[\cdot, \cdot]_A$ satisfies the Jacobi identity on $A$, it suffices to consider triples of elements $x, y, \text{ and } z$ of $A$ where each of $x, y, \text{ and } z$ is equal to either $a_0$ or $b_0$, because of bilinearity. This means that at least two of the elements $x, y, \text{ and } z$ are the same. In this situation, the Jacobi identity can be obtained from (9.6.5), as in the preceding paragraph.

9.7 Nilpotency conditions

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. Remember that if $x \in A$, then $\text{ad } x = \text{ad}_x$ is the homomorphism from $A$ into itself, as a module over $k$, defined by

\begin{equation}
\text{ad}_x(y) = [x, y]_A
\end{equation}
for every \( y \in A \), as in Section 2.4. Let \( n \) be a positive integer, and let \( x_1, \ldots, x_n \) be \( n \) elements of \( A \). Thus \( \text{ad}_{x_j} \) is a module homomorphism from \( A \) into itself for each \( j = 1, \ldots, n \), so that compositions of the \( \text{ad}_{x_j} \)'s are defined as module homomorphisms from \( A \) into itself. If \( y \in A \), then

\[
(9.7.2) \quad (\text{ad}_{x_1} \circ \text{ad}_{x_2} \circ \cdots \circ \text{ad}_{x_n})(y) = [x_1, [x_2, \ldots, [x_n, y]_A, \ldots]_A]_A.
\]

This is an element of the ideal \( A^n \) defined in Section 9.5. If \( A^n = \{0\} \), then (9.7.2) is equal to 0 for every \( y \in A \), so that

\[
(9.7.3) \quad \text{ad}_{x_1} \circ \text{ad}_{x_2} \circ \cdots \circ \text{ad}_{x_n} = 0
\]

as a mapping from \( A \) into itself. Conversely, if (9.7.3) holds for every \( x_1, \ldots, x_n \) in \( A \), then (9.7.2) is equal to 0 for all \( x_1, \ldots, x_n, y \in A \), and which implies that \( A^n = \{0\} \). This corresponds to part of Theorem 2.1 on p32 of [24], and is also mentioned on p12 of [13].

Let \( B \) be an associative algebra over \( k \), where multiplication of \( b, b' \in B \) is expressed as \( bb' \). An element \( b \) of \( B \) is said to be \textit{nilpotent} if \( b^l = 0 \) for some positive integer \( l \). If \( b_1 \) and \( b_2 \) are commuting nilpotent elements of \( B \), then it is easy to see that \( b_1 + b_2 \) is nilpotent as well. More precisely, if \( b_1^{l_1} = 0 \) and \( b_2^{l_2} = 0 \) for some positive integers \( l_1 \) and \( l_2 \), then

\[
(9.7.4) \quad (b_1 + b_2)^{l_1 + l_2 - 1} = 0.
\]

Indeed, the left side of (9.7.4) can be expressed as a sum of terms of the form \( b_1^{j_1} b_2^{j_2} \), where \( j_1 \) and \( j_2 \) are nonnegative integers with \( j_1 + j_2 = l_1 + l_2 - 1 \), which implies that either \( j_1 \geq l_1 \) or \( j_2 \geq l_2 \).

Let \( B_{\text{Lie}} \) be \( B \) as a Lie algebra over \( k \), with respect to the commutator bracket \([b, b']_B = bb' - b'b \) corresponding to multiplication on \( B \). Let \( n \) be a positive integer, and suppose that

\[
(9.7.5) \quad b_1 b_2 \cdots b_n b_{n+1} = 0
\]

for every \( b_1, b_2, \ldots, b_n, b_{n+1} \in B \). This implies that

\[
(9.7.6) \quad [b_1, [b_2, \ldots, [b_n, b_{n+1}]_B \ldots]_B]_B = 0
\]

for every \( b_1, b_2, \ldots, b_n, b_{n+1} \in B \), because the left side of (9.7.6) can be expanded into a sum of products of \( n+1 \) elements of \( B \). It follows that \( B_{n, \text{Lie}}^n = \{0\} \), where \( B_{n, \text{Lie}}^n \) is defined as in Section 9.5. In particular, this means that \( B_{\text{Lie}}^n \) is nilpotent as a Lie algebra.

Let \( A \) be a Lie algebra over \( k \) again. An element \( x \) of \( A \) is said to be \textit{ad-nilpotent} if \( \text{ad}_x \) is nilpotent as an element of the algebra of module homomorphisms from \( A \) into itself, as on p12 of [13]. If \( A^n = \{0\} \) for some positive integer \( n \), then \( (\text{ad}_x)^n = 0 \) as a module homomorphism from \( A \) into itself for every \( x \in A \), as in (9.7.3). In particular, if \( A \) is nilpotent as a Lie algebra, then every element of \( A \) is ad-nilpotent.
Let $B$ be an associative algebra over $k$ again, with corresponding Lie algebra $B_{Lie}$. If $b \in B$, then $\text{ad}_b$ is defined as a module homomorphism from $B$ into itself, by

$$\text{ad}_b(c) = [b, c]_B = b c - c b$$

for every $c \in B$. Equivalently,

$$\text{ad}_b = M_b - \tilde{M}_b,$$

where $M_b$ and $\tilde{M}_b$ are the operators of left and right multiplication by $b$ on $B$, respectively, as in Sections 2.2 and 2.7. Remember that $M_b$ and $\tilde{M}_b$ commute as module homomorphisms from $B$ into itself. If $b^l = 0$ for some positive integer $l$, then $(M_b)^l = M_{b^l} = 0$ and $(\tilde{M}_b)^l = \tilde{M}_{b^l} = 0$, as module homomorphisms from $B$ into itself. This implies that $(\text{ad}_b)^{2l-1} = 0$, as a module homomorphism from $B$ into itself, as in (9.7.4). Thus $b$ is ad-nilpotent as an element of $B_{Lie}$ when $b$ is nilpotent in $B$. This corresponds to the lemma on p12 of [13], and Step 2 on p34 of [24].

### 9.8 Maximal Lie subalgebras

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. If $x \in A$, then

$$\{\alpha x : \alpha \in k\}$$

is a Lie subalgebra of $A$ that contains $x$ as an element. If $x \neq 0$ and $k$ is a field, then (9.8.1) is one-dimensional as a linear subspace of $A$.

Let $B$ be a Lie subalgebra of $A$ that is proper, so that $B \neq A$. As usual, $B$ is said to be maximal as a proper Lie subalgebra of $A$ with respect to inclusion if any proper Lie subalgebra of $A$ that contains $B$ is equal to $B$. If $B_0$ is any proper Lie subalgebra of $A$, then in some situations one can get the existence of a maximal proper Lie subalgebra $B$ of $A$ that contains $B_0$. In particular, if $k$ is a field and $A$ is finite-dimensional as a vector space over $k$, then one can take $B$ to be a Lie subalgebra of $A$ that contains $B_0$ and whose dimension is strictly less than the dimension of $A$ and maximal. One can also get such maximal proper Lie subalgebras of $A$ when Lie subalgebras of $A$ satisfy an ascending chain condition.

Let $B_0$ be a submodule of $A$, as a module over $k$. The normalizer $N_A(B_0)$ of $B_0$ in $A$ is defined to be the set of $x \in A$ such that

$$[x, y]_A \in B_0$$

for every $y \in B_0$. It is easy to see that $N_A(B_0)$ is a submodule of $A$, as a module over $k$, because $B_0$ is a submodule of $A$. One can check that $N_A(B_0)$ is a Lie subalgebra of $A$, using the Jacobi identity. If $B_0$ is a Lie subalgebra of $A$, then

$$B_0 \subseteq N_A(B_0),$$
and in fact $B_0$ is an ideal in $N_A(B_0)$, as a Lie algebra over $k$. In this case, $N_A(B_0)$ may be described as the largest Lie subalgebra of $A$ that contains $A$ as an ideal, as on p7 of [13]. If $B_0 = N_A(B_0)$, then $B_0$ is said to be self-normalizing in $A$. If $B_0$ is a maximal Lie subalgebra of $A$, then it follows that $B_0$ is either self-normalizing in $A$, or $B_0$ is an ideal in $A$.

Let $B$ be an ideal in $A$. If $x \in A$, then it is easy to see that
\[
\{ \alpha x + y : \alpha \in k, y \in B \}
\]
is a Lie subalgebra of $A$. Of course, (9.8.4) contains $B$ and $x$. If $B$ is a maximal proper Lie subalgebra of $A$, and $x \in A \setminus B$, then (9.8.4) is equal to $A$.

Suppose that $A \neq \{0\}$, and that $\{0\}$ is maximal as a proper Lie subalgebra of $A$. Let $x$ be a nonzero element of $A$, so that (9.8.1) is equal to $A$. Observe that
\[
\{ \alpha \in k : \alpha x = 0 \}
\]
is a proper ideal in $k$, because $x \neq 0$. One can check that (9.8.5) is a maximal ideal in $k$ in this situation. This uses the fact that ideals in $k$ correspond to submodules of (9.8.1), which are Lie subalgebras of $A$.

9.9 Nilpotent linear mappings

Let $k$ be a field, and let $V$ be a vector space over $k$ of positive finite dimension. Remember that the space $L(V)$ of linear mappings from $V$ into itself is an associative algebra over $k$ with respect to composition of mappings. In particular, a linear mapping $T$ from $V$ into itself is said to be nilpotent if $T^n = 0$ for some positive integer $n$, which is to say that $T$ is nilpotent as an element of $L(V)$.

As in Section 2.10, we may use $gl(V)$ to denote the space of linear mappings from $V$ into itself as a Lie algebra over $k$, with respect to the corresponding commutator bracket. Let $A$ be a Lie subalgebra of $gl(V)$, and suppose that every element of $A$ is nilpotent as an element of $L(V)$. Under these conditions, it is well known that there is a $v \in V$ such that $v \neq 0$ and
\[
a(v) = 0
\]
for every $a \in A$. This is the theorem stated at the bottom of p12 in [13], which corresponds to Theorem 3.2’ on p33 in [24], as in Step 1 on p34 in [24].

Note that $A$ is a finite-dimensional vector space over $k$. The theorem is proved using induction on the dimension of $A$. Of course, the theorem is trivial when $A = \{0\}$. If the dimension of $A$ is equal to 1, then the theorem reduces to the fact that a nilpotent linear mapping from $V$ into itself has a nontrivial kernel.

Suppose now that $A$ has positive dimension, and that the theorem holds for Lie algebras over $k$ of smaller dimension. Let $B$ be a proper Lie subalgebra of $A$, which has dimension less than the dimension of $A$.

If $x \in gl(V)$, then let $ad_x$ be the linear mapping from $gl(V)$ into itself defined by
\[
ad_x(y) = [x, y]
\]
for every \( y \in \mathfrak{gl}(V) \), as in Section 2.4. If \( x \) is nilpotent as an element of \( \mathcal{L}(V) \), then \( \text{ad}_x \) is nilpotent as a linear mapping from \( \mathfrak{gl}(V) \) into itself, as in Section 9.7. If \( x \in A \), then \( \text{ad}_x \) maps \( A \) into itself, because \( A \) is a Lie subalgebra of \( \mathfrak{gl}(V) \), by hypothesis. More precisely, \( \text{ad}_x \) is nilpotent as a linear mapping from \( A \) into itself, because \( x \) is nilpotent as an element of \( \mathcal{L}(V) \), by hypothesis.

If \( x \in B \), then \( \text{ad}_x \) maps \( B \) into itself, because \( B \) is a Lie subalgebra of \( A \). In particular, \( B \) is a linear subspace of \( A \), so that the quotient \( A/B \) can be defined as a vector space over \( k \). Let \( q \) be the canonical quotient mapping from \( A \) onto \( A/B \). If \( T \) is a linear mapping from \( A \) into itself that maps \( B \) into itself, then \( q \circ T \) is a linear mapping from \( A \) into \( A/B \) whose kernel contains \( B \). This means that for \( a \in A \), \( q(T(a)) \) only depends on \( q(a) \), so that there is a unique linear mapping \( T^{A/B} \) from \( A/B \) into itself such that

\[
T^{A/B} \circ q = q \circ T.
\] (9.9.3)

Note that \( T^{A/B} \) is nilpotent on \( A/B \) when \( T \) is nilpotent on \( A \). Of course, \( A/B \neq \{0\} \), because \( B \neq A \), by hypothesis.

If \( x \in B \), then we can apply this to \( T = \text{ad}_x \), to get a linear mapping \( \text{ad}_x^{A/B} \) from \( A/B \) into itself. More precisely, \( \text{ad}_x^{A/B} \) is nilpotent on \( A/B \), because \( \text{ad}_x \) is nilpotent on \( A \), as before. Let \( C \) be the collection of \( \text{ad}_x^{A/B} \), with \( x \in B \). This is a Lie subalgebra of \( \mathfrak{gl}(A/B) \). The dimension of \( C \), as a vector space over \( k \), is less than or equal to the dimension of \( B \).

Hence the dimension of \( C \) is strictly less than the dimension of \( A \). The induction hypothesis permits us to apply the theorem to \( C \), to get that there is a nonzero element \( q(a_0) \) of \( A/B \), \( a_0 \in A \), such that

\[
\text{ad}_x^{A/B}(q(a_0)) = 0
\] (9.9.4)

for every \( x \in B \). Equivalently, this means that \( q(\text{ad}_x(a_0)) = 0 \) for every \( x \in B \), which is the same as saying that

\[
[x, a_0] = \text{ad}_x(a_0) \in B
\] (9.9.5)

for every \( x \in B \). This shows that \( a_0 \) is an element of the normalizer \( N_A(B) \) of \( B \) in \( A \). Note that \( a_0 \notin B \), because \( q(a_0) \neq 0 \) in \( A/B \).

Let us now take \( B \) to be a maximal proper Lie subalgebra of \( A \). In this case, we get that \( B \) is an ideal in \( A \), as in the previous section. Put

\[
W = \{v \in V : x(v) = 0 \text{ for every } x \in B\},
\] (9.9.6)

which is a linear subspace of \( V \). Using the induction hypothesis again, we get that \( W \neq 0 \). Let us check that

\[
a(W) \subseteq W
\] (9.9.7)

for every \( a \in A \). If \( x \in B \) and \( v \in W \), then

\[
x(a(v)) = a(x(v)) - ([a, x])(v) = 0,
\] (9.9.8)
because \([a,x] \in B\), as before. This implies that \(a(v) \in W\), as desired.

Let \(a_1\) be any element of \(A \setminus B\). The restriction of \(a_1\) to \(W\) is a nilpotent linear mapping from \(W\) into itself, and hence there is a \(v_1 \in W\) such that \(v_1 \neq 0\) and \(a_1(v_1) = 0\). If \(a\) is any element of \(A\), then \(a\) can be expressed as the sum of an element of \(B\) and a scalar multiple of \(a_1\), as in the previous section. It follows that \(a(v_1) = 0\), because \(v_1 \in W\). Thus (9.9.1) holds, as desired.

9.10 Engel’s theorem

Let \(k\) be a field, and let \(V\) be a vector space over \(k\) of positive finite dimension \(n\). A finite sequence \(F = \{V_j\}_{j=0}^{n}\) of \(n\) linear subspaces in \(V\) is said to be a flag in \(V\) if \(V_0 = \{0\}, V_n = V, V_j \subseteq V_{j+1}\) for \(j = 0, \ldots, n-1\), and the dimension of \(V_j\) is equal to \(j\) for each \(j = 0, \ldots, n\). If \(l\) is a nonnegative integer with \(l \leq n\), then let \(U_l(F)\) be the collection of linear mappings \(T\) from \(V\) into itself such that \(T(V_j) \subseteq V_{j-l}\) for each \(j = l, \ldots, n\). This is a subalgebra of the algebra \(L(V)\) of all linear mappings from \(V\) into itself, as an associative algebra over \(k\) with respect to composition of mappings. Note that \(U_0(F)\) contains the identity mapping \(I = I_V\) on \(V\), \(U_n(F) = 0\), and \(U_{l_1}(F) \subseteq U_{l_1}(U_{l_2}(F))\) when \(l_1 \leq l_2\).

More precisely, if \(l_1, l_2\) are nonnegative integers with \(l_1 + l_2 \leq n\), \(T_1 \in U_{l_1}(F)\), and \(T_2 \in U_{l_2}(F)\), then \(T_1 \circ T_2 \in U_{l_1+l_2}(F)\).

In particular, \(U_l(F)\) is an ideal in \(U_0(F)\) for each \(l\). If \(T_1, T_2, \ldots, T_n \in U_1(F)\), then \(T_1 \circ T_2 \circ \cdots \circ T_n = 0\), because the \(n\)-fold composition on the left is an element of \(U_n(F)\). We may consider \(U_l(F)\) as a Lie subalgebra of the Lie algebra \(gl(V)\) of linear mappings from \(V\) into itself with respect to the commutator bracket for each \(l = 0, 1, \ldots, n\). It follows from (9.10.4) that \(U_l(F)\) is nilpotent as a Lie algebra over \(k\), with respect to the commutator bracket, as in Section 9.7.

Let \(A\) be a Lie subalgebra of \(gl(V)\), and suppose that every element of \(A\) is nilpotent as a linear mapping on \(V\), as in the previous section. Under these conditions, it is well known that there is a flag \(F\) in \(V\) such that

\[
A \subseteq U_l(F).
\]

This is the corollary stated on p13 of [13], which corresponds to Theorem 3.2 on p33 of [24]. If \(V \neq \{0\}\), then one can first get a one-dimensional linear subspace \(V_1\) of \(V\) on which the elements of \(A\) vanish, as in the previous section. In order to repeat the process, one can look at the induced linear mappings on the quotient \(V/V_1\).
Now let $A$ be any finite-dimensional Lie algebra over $k$. If every element of $A$ is ad-nilpotent, as in Section 9.7, then it is well known that $A$ is nilpotent as a Lie algebra. This is the theorem stated on the middle of p12 in [13], which corresponds to Theorem 3.1 on p33 of [24]. In the argument on the bottom of p34 of [24], one applies the theorem mentioned in the preceding paragraph to the image of $A$ under the adjoint representation, as a Lie algebra of linear mappings from $A$ into itself. This leads to a flag of linear subspaces of $A$, which are in fact ideals in $A$, and which can be used to show that $A$ is nilpotent as a Lie algebra, as in Section 9.5. Alternatively, one can use the previous theorem to get that the image of $A$ under the adjoint representation is nilpotent as a Lie algebra. This implies that $A$ is nilpotent as a Lie algebra, because the kernel of the adjoint representation is the center $Z(A)$ of $A$, as a Lie algebra.

Let $A$ be a finite-dimensional nilpotent Lie algebra over $k$, and let $B$ be an ideal in $A$ with $B \neq \{0\}$. If $x \in A$, then $\text{ad}_x$ is a linear mapping from $A$ into itself that maps $B$ into itself, because $B$ is an ideal in $A$. Consider the collection $A_B$ of linear mappings from $B$ into itself obtained by restricting $\text{ad}_x$ to $B$ for each $x \in A$. This is a Lie subalgebra of $\text{gl}(B)$, because of the usual properties of the adjoint representation on $A$, as in Section 2.4. Remember that $\text{ad}_x$ is nilpotent as a Lie mapping from $A$ into itself for each $x \in A$, because $A$ is nilpotent as a Lie algebra, as in Section 9.7. This implies that the elements of $A_B$ are nilpotent as linear mappings from $B$ into itself. It follows that there is a $y \in B$ such that $y \neq 0$ and $\text{ad}_x(y) = 0$ for every $x \in A$, as in the previous section. Equivalently, this means that $B \cap Z(A) \neq \{0\}$, as in the lemma on p13 of [13].

### 9.11 Flags and matrices

Let $k$ be a field, let $V$ be a vector space of positive finite dimension $n$, and let $\mathcal{F} = \{V_j\}_{j=0}^n$ be a flag in $V$. If $T$ is an element of the algebra $\mathcal{U}_0(\mathcal{F})$ defined in the previous section, then $T$ induces a linear mapping from $V_j/V_{j-1}$ into itself for each $j = 1, \ldots, n$. This linear mapping corresponds to multiplication by an element $\phi_j(T)$ of $k$, because $V_j/V_{j-1}$ is a one-dimensional vector space over $k$. This defines an algebra homomorphism $\phi_j$ from $\mathcal{U}_0(\mathcal{F})$ into $k$ for each $j = 1, \ldots, n$. Let us consider $k^n$ as a commutative associative algebra over $k$, with respect to coordinatewise addition and multiplication. Thus we get an algebra homomorphism $\phi$ from $\mathcal{U}_0(\mathcal{F})$ into $k^n$, whose $j$th coordinate is the algebra homomorphism $\phi_j$ from $\mathcal{U}_0(\mathcal{F})$ into $k$ just mentioned. The kernel of $\phi$ is the ideal $\mathcal{U}_1(\mathcal{F})$ of $\mathcal{U}_0(\mathcal{F})$ defined in the previous section. In particular, if $T_1, T_2 \in \mathcal{U}_0(\mathcal{F})$, then

\begin{equation}
T_1 \circ T_2 - T_2 \circ T_1 \in \mathcal{U}_1(\mathcal{F}).
\end{equation}
9.11. FLAGS AND MATRICES

Remember that $U_0(F)$ is a Lie subalgebra of $gl(V)$, so that $U_0(F)$ may be considered as a Lie algebra over $k$ with respect to the commutator bracket associated to composition of linear mappings on $V$. Using (9.11.1), we get that

\[(9.11.2) \quad [U_0(F), U_0(F)] \subseteq U_0(F).\]

This implies that $U_0(F)$ is solvable as a Lie algebra, because $U_1(F)$ is nilpotent as a Lie algebra.

Now let $k$ be a commutative ring with a multiplicative identity element, let $n$ be a positive integer, and let $A$ be an associative algebra over $k$, where multiplication of $x, y \in A$ is expressed as $xy$. Remember that the space $M_n(A)$ of $n \times n$ matrices with entries in $A$ is an associative algebra over $k$ with respect to matrix multiplication, as in Section 2.8. If $r$ is a nonnegative integer with $r \leq n$, then let $T_{n,r}(A)$ be the collection of $a = (a_{j,l}) \in M_n(A)$ such that

\[(9.11.3) \quad a_{j,l} = 0\]

when $l \leq j + r - 1$. Equivalently, this means that $a_{j,l}$ may be nonzero only when $l \geq j + r$. Thus $T_{n,0}(A)$ consists of upper-triangular matrices, $T_{n,1}(A)$ consists of strictly upper-triangular matrices, and $T_{n,n}(A) = \{0\}$. Clearly

\[(9.11.4) \quad T_{n,r_1}(A) \subseteq T_{n,r_2}(A)\]

when $r_1 \leq r_2$. If $a \in T_{n,r_1}(A)$ and $b \in T_{n,r_2}(A)$ for some nonnegative integers $r_1, r_2$ with $r_1 + r_2 \leq n$, then

\[(9.11.5) \quad a b \in T_{n,r_1 + r_2}(A).\]

In particular, $T_{n,0}(A)$ is a subalgebra of $M_n(A)$, $T_{n,r}(A)$ is an ideal in $T_{n,0}(A)$ for each $0 \leq r \leq n$, and the product of $n$ elements of $T_{n,1}(A)$ is equal to $0$.

If $j$ is a positive integer with $j \leq n$ and $a \in T_{n,0}(A)$, then put $\psi_j(a) = a_{j,j}$, which defines an algebra homomorphism from $T_{n,0}(A)$ onto $A$. Let $\psi$ be the mapping from $T_{n,0}(A)$ into $A^n$ such that the $j$th coordinate of $\psi(a)$ is equal to $\psi_j(a)$ for every $j = 1, \ldots, n$ and $a \in T_{n,0}(A)$. This defines an algebra homomorphism from $T_{n,0}(A)$ onto $A^n$, where $A^n$ is considered as an associative algebra over $k$ with respect to coordinatewise addition and multiplication. The kernel of $\psi$ is equal to $T_{n,1}(A)$. If $A$ is a commutative algebra over $k$, then

\[(9.11.6) \quad a b - b a \in T_{n,1}(A)\]

for every $a, b \in T_{n,0}(A)$.

Remember that $gl_n(A)$ is the same as $M_n(A)$, but considered as a Lie algebra over $k$ with respect to the commutator bracket. Similarly, we may use $t_{n,r}(A)$ for $T_{n,r}(A)$, considered as a Lie subalgebra of $gl_n(A)$. As in Section 9.7, $t_{n,1}(A)$ is nilpotent as a Lie algebra, because the product of $n$ elements of $T_{n,1}(A)$ is equal to $0$. If $A$ is commutative, then

\[(9.11.7) \quad [t_{n,0}(A), t_{n,0}(A)] \subseteq t_{n,1}(A),\]
by (9.11.6). This implies that $t_{n,0}(A)$ is solvable as a Lie algebra, because $t_{n,1}(A)$ is nilpotent.

Let $k$ be a field again, and let $V$ be a vector space over $k$ of dimension $n \in \mathbb{Z}_+$. If $v_1, \ldots, v_n$ is a basis for $V$, then we can get a flag $F = \{V_j\}_{j=0}^n$ in $V$ by taking $V_j$ to be the linear span of $v_1, \ldots, v_j$ for each $j = 1, \ldots, n$. Of course, every flag in $V$ corresponds to a basis for $V$ in this way. Using this basis for $V$, elements of $M_n(k)$ correspond to linear mappings from $V$ into itself, as in Section 2.10. Similarly, $T_{n,r}(k)$ corresponds to $U_r(F)$ for each $r = 0, 1, \ldots, n$.

9.12 A useful lemma

Let $k$ be a field of characteristic 0, let $(A, [\cdot, \cdot])$ be a Lie algebra over $k$, and let $B$ be an ideal in $A$. Also let $V$ be a finite-dimensional vector space over $k$, and suppose that $V$ is a module over $A$, as a Lie algebra over $k$. Let $v$ be a nonzero element of $V$, and suppose that $\chi$ is a mapping from $B$ into $k$ such that

\begin{equation}
(9.12.1) \quad b \cdot v = \chi(b) v
\end{equation}

for every $b \in B$. If $a \in A$ and $b \in B$, then $[a, b] \in B$, and in fact

\begin{equation}
(9.12.2) \quad \chi([a, b]) = 0.
\end{equation}

This is the Main Lemma stated on p36 of [24], which corresponds to part of the proof of Step (3) on p16 of [13].

Equivalently, let $\rho$ be the representation of $A$ on $V$, which makes $V$ into a module over $A$ as a Lie algebra over $k$. The hypothesis (9.12.1) says that for each $b \in B$, $v$ is an eigenvector of $\rho_b$, with eigenvalue $\chi(b)$. The conclusion (9.12.2) says that

\begin{equation}
(9.12.3) \quad (\rho([a, b]) \cdot v = \rho([a, b])v = 0
\end{equation}

for every $a \in A$ and $b \in B$. One could also reduce to the case where $A$ is a Lie subalgebra of $gl(V)$, by considering the Lie algebra of linear mappings from $V$ into itself of the form $\rho_a$ for some $a \in A$. In particular, the discussion in [13] is given in this setting.

Let $a \in A$ be given, and put $V_0 = \{0\}$. If $j$ is a positive integer, then let $V_j$ be the linear span of

\begin{equation}
(9.12.4) \quad v, \rho_a(v), \ldots, (\rho_a)^{j-1}(v)
\end{equation}

in $V$. Thus $V_j$ is a linear subspace of $V$ for every $j \geq 0$, with $V_j \subseteq V_{j+1}$. Let $n$ be the smallest positive integer such that

\begin{equation}
(9.12.5) \quad V_n = V_{n+1}.
\end{equation}

This uses the finite-dimensionality of $V$, to get that this condition holds for some positive integer. Note that $V_n$ has dimension equal to $n$ as a vector space over $k$, and that

\begin{equation}
(9.12.6) \quad \rho_a(V_n) \subseteq V_{n+1} = V_n.
\end{equation}
In particular, \( n \) is less than or equal to the dimension of \( V \).

If \( b \in B \), then we would like to show that

\[
\rho_b(\rho_n(v)^j) = \chi(b) \rho_n(v)^j \mod V_j
\]

for each \( j \geq 0 \), using induction on \( j \). This is the same as (9.12.1) when \( j = 0 \).

If \( j \geq 1 \), then

\[
\rho_b(\rho_n(v)^j) = \rho_b\rho_n(v)^j - \rho_n(v)^j(\rho_a(v))
\]

\[
= \rho_a\rho_b(v)^j - \rho_{[a,b]}(v)^j.
\]

Of course,

\[
\rho_b(\rho_n(v)^{j-1}) = \chi(b) \rho_n(v)^{j-1} \mod V_{j-1},
\]

by the induction hypothesis. This implies that

\[
\rho_a(\rho_b(\rho_n(v)^{j-1})) = \chi(b) \rho_n(v)^j \mod V_j,
\]

because \( \rho_a(V_{j-1}) \subseteq V_j \) by construction. Similarly,

\[
\rho_{[a,b]}(\rho_n(v)^{j-1}) = \chi([a,b]) \rho_n(v)^{j-1} \mod V_{j-1},
\]

by the induction hypothesis, because \( [a,b] \in B \). It follows that the left side of

(9.12.11) is an element of \( V_j \). Combining this with (9.12.8) and (9.12.10), we
get that (9.12.7) holds, as desired.

In particular, (9.12.7) implies that \( \rho_b \) maps \( V_n \) into itself when \( b \in B \). Observe that

\[
\tr_{V_n} \rho_b = n \cdot \chi(b)
\]

for every \( b \in B \), by (9.12.7), where more precisely the left side is the trace of the restriction of \( \rho_b \) to \( V_n \). We also have that

\[
\tr_{V_n} \rho_{[a,b]} = \tr_{V_n} [\rho_n, \rho_b] = 0
\]

for every \( b \in B \), using the fact that \( \rho_a \) and \( \rho_b \) both map \( V_n \) into itself in the
second step. Thus

\[
n \cdot \chi([a,b]) = 0
\]

for every \( b \in B \), by (9.12.12) applied to \( [a,b] \). This implies (9.12.2), because \( k \) is
supposed to have characteristic 0. Note that this also works when \( k \) has positive characteristic and the dimension of \( V \) is strictly less than the characteristic of \( k \), because \( n \) is less than or equal to the dimension of \( V \). This is related to
Exercise 2 on p20 of [13].
9.13 Lie’s theorem

Let $k$ be an algebraically closed field of characteristic 0, let $(A, [\cdot, \cdot])$ be a solvable Lie algebra over $k$, and let $\rho$ be a representation of $A$ as a Lie algebra on a finite-dimensional vector space $V$ over $k$. If $V \neq \{0\}$, then there exists a $v \in V$ such that $v \neq 0$ and $v$ is an eigenvector for $\rho_a$ for every $a \in A$. This corresponds to Theorem 5.1’ on p36 of [24], and the theorem on p15 of [13]. As before, one can reduce to the case where $A$ is a Lie subalgebra of $gl(V)$, by considering the Lie algebra of linear mappings from $V$ into itself of the form $\rho_a$ for some $a \in A$. The theorem on p15 of [13] is stated in this way, so that $A$ is finite-dimensional as a vector space over $k$ in particular. The finite-dimensionality of $A$ is implicit in Theorem 5.1’ in [24], as mentioned at the beginning of Chapter 5 in [24]. Let us suppose now that $A$ is finite-dimensional as a vector space over $k$ too.

The proof uses induction on the dimension of $A$, as a vector space over $k$. Of course, if $A = \{0\}$, then the statement is trivial. Suppose now that $A \neq \{0\}$, and note that $[A, A] \neq A$, because $A$ is solvable. Let $B$ be a linear subspace of $V$ of codimension 1 that contains $[A, A]$, which implies that $B$ is an ideal in $A$. The induction hypothesis implies that there is a $v \in V$ with $v \neq 0$ and a mapping $\chi$ from $B$ into $k$ such that

\[(9.13.1)\quad \rho_a(v) = \chi(b) v\]

for every $b \in B$. Put

\[(9.13.2)\quad W = \{w \in V : \rho_a(w) = \chi(b) w \text{ for every } b \in B\},\]

which is a linear subspace of $V$ with $v \in W$, so that $W \neq \{0\}$. We would like to verify that

\[(9.13.3)\quad \rho_a(W) \subseteq W\]

for every $a \in A$. If $a \in A$, $b \in B$, and $w \in W$, then

\[(9.13.4)\quad \rho_a(\rho_b(w)) = \rho_b(\rho_a(w)) - \rho_{[a, b]}(w) = \chi(b) \rho_a(w) - \chi([a, b]) w,\]

using the fact that $[a, b] \in B$ in the second step. Combining this with (9.12.2), we obtain that

\[(9.13.5)\quad \rho_b(\rho_a(w)) = \chi(b) \rho_a(w),\]

as desired.

Let $a_0$ be any element of $A$ not in $B$. Because $k$ is algebraically closed, there is a $w_0 \in W$ such that $w_0 \neq 0$ and $w_0$ is an eigenvector for $\rho_{a_0}$. If $a$ is any element of $A$, then $a$ can be expressed as the sum of a multiple of $a_0$ and an element $b$ of $B$, because $B$ has codimension 1 in $A$. It follows that $w_0$ is an eigenvector for $\rho_a$, as desired, because $w_0$ is an eigenvector for $\rho_{a_0}$ by definition of $W$. Note that this also works when $k$ has positive characteristic strictly larger than the dimension of $V$, as in Exercise 2 on p20 of [13].

Under these conditions, Lie’s theorem states that there is a flag $F = \{V_j\}_{j=0}^n$ in $V$ such that $\rho_a(V_j) \subseteq V_j$ for every $a \in A$ and $j = 0, 1, \ldots, n$. This is Theorem
5.1 on p36 of [24], which corresponds to Corollary A on p16 of [13]. More precisely, one can get \( V_1 \) as in the previous paragraphs. One can repeat the process, by considering the induced linear mappings on \( V/V_1 \).

In particular, if \( A \) is a solvable Lie algebra over \( k \) that is finite-dimensional as a vector space over \( k \), then there is a flag in \( A \) consisting of ideals in \( A \). This is Corollary 5.2 on p37 of [24], and Corollary B on p16 of [13]. This follows from the statement in the preceding paragraph, applied to the adjoint representation of \( A \).

**9.14 Structure constants**

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( n \) be a positive integer. Remember that the space \( k^n \) of \( n \)-tuples of elements of \( k \) is a \((free)\) module over \( k \) with respect to coordinatewise addition and scalar multiplication. Let \( u_1, \ldots, u_n \) be the “standard basis” elements of \( k^n \), so that the \( j \)th coordinate of \( u_l \) is equal to 1 when \( j = l \) and to 0 otherwise. If \( x = (x_1, \ldots, x_n) \in k^n \), then

\[
x = \sum_{l=1}^{n} x_l u_l.
\]

(9.14.1)

Of course, if \( k \) is a field and \( V \) is an \( n \)-dimensional vector space over \( k \), then \( V \) can be identified with \( k^n \) by choosing a basis for \( V \).

Let \( [\cdot, \cdot]_{k^n} \) be a mapping from \( k^n \times k^n \) into \( k^n \) that is bilinear over \( k \). We can express \( [u_j, u_l]_{k^n} \) as

\[
[u_j, u_l]_{k^n} = \sum_{r=1}^{n} c_{j,l}^{r} u_r
\]

(9.14.2)

for each \( j, l = 1, \ldots, n \), where \( c_{j,l}^{r} \) are elements of \( k \) for every \( j, l, r = 1, \ldots, n \). This implies that

\[
([x, y]_{k^n})_r = \sum_{j=1}^{n} \sum_{l=1}^{n} c_{j,l}^{r} x_j y_l
\]

(9.14.3)

for every \( x, y \in k^n \) and \( r = 1, \ldots, n \), where the left side is the \( r \)th coordinate of \( [x, y]_{k^n} \). More precisely, this uses (9.14.1) and the bilinearity of \( [\cdot, \cdot]_{k^n} \) over \( k \). Conversely, if \( c_{j,l}^{r} \in k \) for every \( j, l, r = 1, \ldots, n \), then (9.14.3) defines a mapping from \( k^n \times k^n \) into \( k^n \) that is bilinear over \( k \).

Clearly

\[
[x, y]_{k^n} = -[y, x]_{k^n}
\]

(9.14.4)

for every \( x, y \in k^n \) if and only if

\[
c_{j,l}^{r} = -c_{l,j}^{r}
\]

(9.14.5)

for every \( j, l, r = 1, \ldots, n \). One can check that

\[
[x, x]_{k^n} = 0
\]

(9.14.6)
for every \( x \in k^n \) if and only if (9.14.5) holds and

\[
(9.14.7) \quad c_{j,j}^r = 0
\]

for every \( j, r = 1, \ldots, n \). If \( 1 + 1 \) has a multiplicative inverse in \( k \), then (9.14.4) implies (9.14.6), and (9.14.5) implies (9.14.7), as usual. The Jacobi identity for \([\cdot, \cdot]_k^n\) holds if and only if

\[
(9.14.8) \quad \sum_{h=1}^{n} (c_{j,l}^h c_{h,m}^r + c_{l,m}^h c_{h,j}^r + c_{m,j}^h c_{h,l}^r) = 0
\]

for every \( j, l, m, r = 1, \ldots, n \), as on p5 of [13]. Let us suppose from now on in this section that the \( c_{j,i}^r \)'s satisfy these conditions, so that \([\cdot, \cdot]_k^n\) defines a Lie bracket on \( k^n \).

Let \( A \) be a commutative associative algebra over \( k \). Note that the space \( A^n \) of \( n \)-tuples of elements of \( A \) is a module over \( k \) with respect to coordinatewise addition and scalar multiplication. If \( a, b \in A^n \), then define \([a, b]_{A^n}\) as an element of \( A^n \) by

\[
(9.14.9) \quad ([a, b]_{A^n})_r = \sum_{j=1}^{n} \sum_{l=1}^{n} c_{j,l}^r a_j b_l
\]

for each \( r = 1, \ldots, n \), where the left side is the \( r \)th coordinate of \([a, b]_{A^n}\). The conditions on the \( c_{j,i}^r \)'s in the preceding paragraph imply that \( A^n \) is a Lie algebra over \( k \) with respect to (9.14.9). If \( k^n \) is solvable or nilpotent as a Lie algebra with respect to \([\cdot, \cdot]_k^n\), then one can check that \( A^n \) has the same property with respect to \([\cdot, \cdot]_{A^n}\).

Suppose that \( A \) has a multiplicative identity element \( e \). In this case, \( A^n \) may be considered as a module over \( A \) with respect to coordinatewise addition and scalar multiplication, and as a Lie algebra over \( A \) with respect to (9.14.9). Of course,

\[
(9.14.10) \quad t \mapsto t e
\]

defines a ring homomorphism from \( k \) into \( A \), which leads to a Lie algebra homomorphism from \( k^n \) into \( A^n \), as Lie algebras over \( k \). Suppose that (9.14.10) is injective, which implies that the corresponding Lie algebra homomorphism from \( k^n \) into \( A^n \) is injective. If \( A^n \) is solvable or nilpotent as a Lie algebra, then it follows that \( k^n \) has the same property.

**9.15 Another corollary**

Let \( k \) be a commutative ring with a multiplicative identity element, and let \((A, [\cdot, \cdot])\) be a Lie algebra over \( k \). If \( x \in A \), then we may use \( \text{ad}_A x = \text{ad}_{A,x} \) to denote the usual mapping

\[
(9.15.1) \quad (\text{ad}_A x)(y) = \text{ad}_{A,x}(y) = [x, y]
\]
9.15. ANOTHER COROLLARY

from $A$ into itself. Similarly, if $B$ is a Lie subalgebra of $A$ and $x \in B$, then $\text{ad}_B x = \text{ad}_B x$ is a module homomorphism from $B$ into itself. In this situation, $\text{ad}_B x$ is the same as the restriction of (9.15.1) to $y \in B$.

Suppose that $k$ is an algebraically closed field of characteristic 0, and that $(A,[\cdot,\cdot])$ be a solvable Lie algebra over $k$ that is finite-dimensional as a vector space over $k$. Under these conditions, $[A,A]$ is nilpotent as a Lie algebra over $k$. This is Corollary C on p16 of [13], and part of Corollary 5.3 on p37 of [24]. Remember that there is a flag of ideals in $A$, as a consequence of Lie’s theorem in Section 9.13. If $x \in A$, then $\text{ad}_{A,x}$ maps these ideals into themselves. If $x \in [A,A]$, then $\text{ad}_{A,x}$ maps the nonzero ideals in the flag into the next smaller one, as in Section 9.11. This implies that $\text{ad}_{A,x}$ is nilpotent as a mapping from $A$ into itself. It follows that $\text{ad}_{[A,A],x}$ is nilpotent as a mapping from $[A,A]$ into itself, because this mapping is the same as the restriction of $\text{ad}_{A,x}$ to $[A,A]$, as in the preceding paragraph. This implies that $[A,A]$ is nilpotent as a Lie algebra, as in Section 9.10.

Alternatively, one can use the same type of argument to get that the image of $[A,A]$ under the adjoint representation of $A$ is nilpotent as a Lie algebra over $k$. One can use this to get that $[A,A]$ is nilpotent as a Lie algebra, because the kernel of the adjoint representation of $A$ is the center of $A$.

Suppose now that $k$ is a field of characteristic 0, and that $A$ is a finite-dimensional solvable Lie algebra over $k$. Corollary 5.3 on p37 of [24] states that $[A,A]$ is still nilpotent as a Lie algebra over $k$, without asking $k$ to be algebraically closed. To see this, let $k_1$ be an algebraically closed field that contains $k$. The statement is trivial when $A = \{0\}$, and so we may suppose that the dimension $n$ of $A$ as a vector space over $k$ is positive. Thus $A$ is isomorphic to $k^n$ as a vector space over $k$, and we may as well suppose that $A = k^n$ with some Lie bracket. This leads to a Lie bracket on $k^n_1$, as in the previous section. If $k^n$ is solvable as a Lie algebra, then $k^n_1$ is solvable as a Lie algebra too, as before. This implies that $[k^n_1,k^n_1]$ is nilpotent as a Lie algebra, by the earlier arguments for algebraically closed fields. Note that $[k^n,k^n]$ may be considered as a Lie subalgebra of $[k^n_1,k^n_1]$, as a Lie algebra over $k$. It follows that $[k^n,k^n]$ is solvable as a Lie algebra over $k$, as desired.
Chapter 10

Matrices and traces

10.1 Some remarks about $\text{gl}_n(k)$

Let $k$ be a commutative ring with a multiplicative identity element, and let $n$ be a positive integer. The space $k^n$ of $n$-tuples of elements of $k$ is a (free) module over $k$ with respect to coordinatewise addition and scalar multiplication, as usual. If $a = (a_{j,l})$ is an $n \times n$ matrix with entries in $k$ and $x \in k^n$, then $T_a(x)$ is defined as the element of $k^n$ whose $j$th coordinate is given by

$$ (T_a(x))_j = \sum_{l=1}^{n} a_{j,l} x_l $$

for each $j = 1, \ldots, n$. This defines a module homomorphism from $k^n$ into itself, and $a \mapsto T_a$ is an algebra isomorphism from the algebra $M_n(k)$ of $n \times n$ matrices with entries in $k$ with respect to matrix multiplication onto the algebra $\text{Hom}_k(k^n, k^n)$ of module homomorphisms from $k^n$ into itself with respect to composition of mappings.

Let $u_1, \ldots, u_n$ be the $n$ “standard basis” elements of $k^n$, so that the $l$th coordinate of $u_r$ is equal to 1 when $l = r$, and to 0 otherwise. Thus

$$ x = \sum_{r=1}^{n} x_r u_r $$

for every $x = (x_1, \ldots, x_n) \in k^n$. If $a \in M_n(k)$, then

$$ (T_a(u_r))_j = a_{j,r} $$

for every $j, r = 1, \ldots, n$.

Similarly, if $h, m \in \{1, \ldots, n\}$, then let $e_{h,m}$ be the element of $M_n(k)$ whose $(h, m)$ entry is equal to 1, and all of whose other entries are equal to 0. If $a = (a_{j,l})$ is any element of $M_n(k)$, then $a$ can be expressed as

$$ a = \sum_{h=1}^{n} \sum_{m=1}^{n} a_{h,m} e_{h,m}. $$
10.1. SOME REMARKS ABOUT $GL_N(K)$

It is sometimes convenient to let $T_{h,m}$ be the module homomorphism from $k^n$ into itself associated to $e_{h,m}$ as in (10.1.1) for each $h, m = 1, \ldots, n$, so that

(10.1.5) \[ T_{h,m} = T_{e_{h,m}}. \]

If $a = (a_{j,l}) \in M_n(k)$, then $T_a$ can be expressed as

(10.1.6) \[ T_a = \sum_{h=1}^{n} \sum_{m=1}^{n} a_{h,m} T_{h,m}. \]

If $q, r \in \{1, \ldots, n\}$, then let $\delta_{q,r} \in k$ be equal to 1 when $q = r$, and to 0 otherwise, as usual. Observe that

(10.1.7) \[ T_{h,m}(u_r) = \delta_{m,r} u_h \]
for every $h, m, r = 1, \ldots, n$. We also have that

(10.1.8) \[ e_{h,m} e_{q,r} = \delta_{m,q} e_{h,r} \]
for every $h, m, q, r = 1, \ldots, n$. Equivalently,

(10.1.9) \[ T_{h,m} \circ T_{q,r} = \delta_{m,q} T_{h,r} \]
for every $h, m, q, r = 1, \ldots, n$. It follows from (10.1.8) that

(10.1.10) \[ [e_{h,m}, e_{q,r}] = e_{h,m} e_{q,r} - e_{q,r} e_{h,m} = \delta_{m,q} e_{h,r} - \delta_{r,h} e_{q,m} \]
for every $h, m, q, r = 1, \ldots, n$. In particular, if $h \neq r$ and $m \neq q$, then

(10.1.11) \[ [e_{h,m}, e_{q,r}] = 0, \]
because each of the two terms on the right side of (10.1.10) is equal to 0. Otherwise,

(10.1.12) \[ [e_{h,m}, e_{m,r}] = e_{h,r} \]
when $h \neq r$, and

(10.1.13) \[ [e_{h,m}, e_{q,h}] = -e_{q,m} \]
when $m \neq q$. Of course, these two cases are equivalent, because of the antisymmetry of the commutator bracket. Similarly,

(10.1.14) \[ [e_{h,m}, e_{m,k}] = e_{h,k} - e_{m,m} \]
for every $h, m = 1, \ldots, n$.

Remember that $gl_n(k)$ is the same as $M_n(k)$ as a module over $k$, but considered as a Lie algebra over $k$ with respect to the corresponding commutator bracket, as in Section 2.9. Similarly, $sl_n(k)$ is the ideal in $gl_n(k)$ consisting of matrices with trace 0, as before. In fact,

(10.1.15) \[ [gl_n(k), gl_n(k)] = sl_n(k), \]
where the left side is as defined in Section 9.2. More precisely, the inclusion of the left side of (10.1.15) in the right side follows from basic properties of the trace, as in Section 2.9. The opposite inclusion can be obtained from (10.1.12) and (10.1.14). This corresponds to Exercise 2 on p9 of [13]. One can also verify that
\[ [gl_n(k), sl_n(k)] = sl_n(k), \]  
(10.1.16) using the same argument.

10.2 Some basic properties of \( sl_2(k) \)

Let \( k \) be a commutative ring with a multiplicative identity element, and remember that \( sl_2(k) \) is the space of \( 2 \times 2 \) matrices with entries in \( k \) and trace 0. This is a Lie algebra over \( k \) with respect to the usual commutator bracket. Consider the elements of \( sl_2(k) \) given by
\[
\begin{align*}
x & = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, & y & = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, & h & = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\end{align*}
\]
(10.2.1)
It is easy to see that every element of \( sl_2(k) \) can be expressed in a unique way as a linear combination of \( x \), \( y \), and \( h \) with coefficients in \( k \). Thus \( sl_2(k) \) is isomorphic to the free module \( k^3 \) of rank 3 over \( k \), as a module over \( k \).

Equivalently, using the notation in the previous section, with \( n = 2 \), we have that
\[
\begin{align*}
x & = e_{1,2}, & y & = e_{2,1}, & h & = e_{1,1} - e_{2,2}.
\end{align*}
\]
(10.2.2)
One can check that
\[
\begin{align*}
[x, y] & = h, & [h, x] & = 2 \cdot x, & [h, y] & = -2 \cdot y,
\end{align*}
\]
(10.2.3)
as on p6 of [13]. Here \( 2 \cdot a = a + a \) for each \( a \in sl_2(k) \), as usual. If \( 1 + 1 = 0 \) in \( k \), then \( h \) is the same as the identity matrix, so that
\[
\begin{align*}
[h, x] & = [h, y] = 0,
\end{align*}
\]
(10.2.4)
as in (10.2.3). In this case, it follows that \( sl_2(k) \) is nilpotent as a Lie algebra over \( k \), as in Exercise 3 on p14 of [13].

Suppose for the moment that \( k \) is a field with characteristic 2. Elements of \( sl_2(k) \) correspond to linear mappings from \( k^2 \) into itself, as a two-dimensional vector space over \( k \), as before. It is well known and not difficult to check that the linear mappings on \( k^2 \) corresponding to elements of \( sl_2(k) \) do not have a (nonzero) simultaneous eigenvector. This shows that the results discussed in Section 9.13 can fail in positive characteristic, as mentioned on p37 of [24].

If \( 1 + 1 \) has a multiplicative inverse in \( k \), then we get that
\[
\begin{align*}
[sl_2(k), sl_2(k)] & = sl_2(k),
\end{align*}
\]
(10.2.5)
where the left side is as defined in Section 9.2. This corresponds to part of Exercise 9 on p5 of [13]. Similarly, if \( 2^j \cdot 1 = 0 \) in \( k \) for some \( j \in \mathbb{Z}_+ \), then one
can check that \( sl_2(k) \) is nilpotent as a Lie algebra over \( k \). However, if for each \( j \in \mathbb{Z}_+ \), \( 2^j \cdot 1 \neq 0 \) in \( k \), then \( sl_2(k) \) is not solvable as a Lie algebra over \( k \).

If \( n \) is any positive integer with \( n \geq 3 \), then

\[
[sl_n(k), sl_n(k)] = sl_n(k).
\]

This corresponds to part of Exercise 9 on p5 of [13] again. Of course, the left side of (10.2.6) is contained in the right side, as in (10.1.15). To get the opposite inclusion, one can verify that

\[
e_{h,r} \in [sl_n(k), sl_n(k)]
\]

for every \( h, r \in \{1, \ldots, n\} \) with \( h \neq r \), using (10.1.12) with \( m \neq h, r \). We also have that

\[
e_{h,h} - e_{m,m} \in [sl_n(k), sl_n(k)]
\]

for every \( h, m = 1, \ldots, n \) with \( h \neq m \), by (10.1.14).

10.3 Scalar and diagonal matrices

Let \( k \) be a commutative ring with a multiplicative identity element, let \( n \) be a positive integer, and let \( A \) be an associative algebra over \( k \), where multiplication of \( x, y \in A \) is expressed as \( xy \). As usual, an \( n \times n \) matrix \( a = (a_{j,l}) \) with entries in \( A \) is said to be a diagonal matrix if \( a_{j,l} = 0 \) when \( j \neq l \). Let \( D_n(A) \) be the space of these diagonal matrices, which is a subalgebra of the algebra \( M_n(A) \) of \( n \times n \) matrices with entries in \( A \), as an associative algebra over \( k \) with respect to matrix multiplication. If multiplication on \( A \) is commutative, then matrix multiplication is commutative on \( D_n(A) \).

An element \( a \) of \( D_n(A) \) is said to be a scalar matrix if the diagonal entries \( a_{j,j} \) of \( a \) are all equal to each other. Let \( S_n(A) \) be the space of these scalar matrices, which is a subalgebra of \( D_n(A) \). If multiplication on \( A \) is commutative, then the elements of \( S_n(A) \) commute with all other elements of \( M_n(A) \), with respect to matrix multiplication.

As before, \( gl_n(A) \) is the same as \( M_n(A) \), but considered as a Lie algebra over \( k \) with respect to the corresponding commutator bracket. Similarly, let \( d_n(A) \) and \( s_n(A) \) be the same as \( D_n(A) \) and \( S_n(A) \), respectively, considered as Lie subalgebras of \( gl_n(A) \).

Let us now simply take \( A = k \). Remember that the identity matrix \( I \) in \( M_n(k) \) is the diagonal matrix with diagonal entries equal to the multiplicative identity element 1 in \( k \), which is the multiplicative identity element in \( M_n(k) \). The scalar matrices in \( M_n(k) \) are the same as scalar multiples of \( I \) by elements of \( k \).

Using the notation in Section 10.1, we have that

\[
e_{h,m} \in sl_n(k)
\]

for every \( h, m = 1, \ldots, n \) with \( h \neq m \), and

\[
e_{h,h} - e_{m,m} \in sl_n(k)
\]
for every \( h, m = 1, \ldots, n \). As on p2 of [13], we may consider the \( e_{h,m} \)'s with \( h \neq m \), together with the matrices \( e_{h,h} - e_{h+1,h+1} \) for \( h = 1, \ldots, n - 1 \) when \( n \geq 2 \), as the “standard basis elements” of \( \mathfrak{sl}_n(k) \). One can check that every element of \( \mathfrak{sl}_n(k) \) can be expressed in a unique way as a linear combination of these standard basis elements with coefficients in \( k \). Note that there are \((n^2 - n) + (n - 1) = n^2 - 1 \) of these standard basis elements in \( \mathfrak{sl}_n(k) \). Thus \( \mathfrak{sl}_n(k) \) is isomorphic to the free module \( k^{n^2-1} \) of rank \( n^2 - 1 \) over \( k \), as a module over \( k \).

Of course, every element of \( \mathfrak{gl}_n(k) \) can be expressed in a unique way as a linear combination of the \( e_{h,m} \)'s, \( h, m = 1, \ldots, n \), with coefficients in \( k \), so that \( \mathfrak{gl}_n(k) \) is isomorphic to the free module \( k^n \) of rank \( n \) over \( k \), as a module over \( k \). Alternatively, one can verify that every element of \( \mathfrak{gl}_n(k) \) can be expressed in a unique way as a linear combination of the standard basis elements for \( \mathfrak{sl}_n(k) \) mentioned in the preceding paragraph together with \( e_{1,1} \), with coefficients in \( k \). In particular, every element of \( \mathfrak{gl}_n(k) \) can be expressed in a unique way as the sum of an element of \( \mathfrak{sl}_n(k) \) and a multiple of \( e_{1,1} \) by an element of \( k \), so that \( \mathfrak{gl}_n(k) \) is isomorphic to the direct sum of \( \mathfrak{sl}_n(k) \) and \( k \), as modules over \( k \). Indeed, if \( a \in \mathfrak{gl}_n(k) \), then

\[
(a - (\operatorname{tr} a) e_{1,1}) \in \mathfrak{sl}_n(k),
\]

because \( \operatorname{tr} e_{1,1} = 1 \). Thus \( a \) can be expressed as the sum of (10.3.3) and \( (\operatorname{tr} a) e_{1,1} \), and one can use the trace again to see that this is unique.

Note that \( \operatorname{tr} I = n \cdot 1 \), as an element of \( k \). If \( n \cdot 1 = 0 \) in \( k \), then \( n \times n \) scalar matrices with entries in \( k \) have trace equal to 0. Suppose for the moment that \( n \cdot 1 \) has a multiplicative inverse in \( k \). If \( a \in \mathfrak{gl}_n(k) \), then

\[
(a - \frac{\operatorname{tr} a}{n \cdot 1} I) \in \mathfrak{sl}_n(k),
\]

and \( \frac{\operatorname{tr} a}{n \cdot 1} \) is the unique element of \( k \) with this property. This implies that every element of \( \mathfrak{gl}_n(k) \) can be expressed in a unique way as the sum of elements of \( \mathfrak{sl}_n(k) \) and \( s_n(k) \), as in Exercise 7 on p5 in [13].

### 10.4 Centrality in \( \mathfrak{gl}_n(k) \), \( \mathfrak{sl}_n(k) \)

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( n \) be a positive integer. Remember that \( a = (a_{j,j}) \in \mathfrak{gl}_n(k) \) can be expressed as

\[
a = \sum_{h=1}^{n} \sum_{m=1}^{n} a_{h,m} e_{h,m},
\]

as in (10.1.4). If \( q, r \in \{1, \ldots, n\} \), then

\[
[a, e_{q,r}] = \sum_{h=1}^{n} \sum_{m=1}^{n} a_{h,m} [e_{h,m}, e_{q,r}]
= \sum_{h \neq r} a_{h,q} e_{h,r} - \sum_{m \neq q} a_{r,m} e_{q,m} + a_{r,q} (e_{r,r} - e_{q,q}),
\]

(10.4.2)
by (10.1.11), (10.1.12) with \( m = q \), (10.1.13) with \( h = r \), and (10.1.14) with \( h = r \) and \( m = q \). More precisely, the two sums on the right side of (10.4.2) are taken over \( h, m = 1, \ldots, n \) with \( h \neq r \) and \( m \neq q \), respectively. If \( q \neq r \), then we get that

\[
[a, e_{q,r}] = \sum_{h \neq q} a_{h,r} e_{h,r} - \sum_{m \neq r} a_{r,m} e_{q,m} + (a_{q,q} - a_{r,r}) e_{q,r} + a_{r,q} (e_{r,r} - e_{q,q}),
\]

(10.4.3)

where the sums are taken over \( h, m = 1, \ldots, n \) with \( h \neq q, r \), respectively.

Suppose that \([a, e_{q,r}] = 0\) for every \( q, r \in \{1, \ldots, n\} \) with \( q \neq r \). In this case, one can use (10.4.3) to get that \( a \) is a diagonal matrix whose diagonal entries are equal to each other, so that \( a \) is a scalar matrix. In particular, the center of \( gl_n(k) \) as a Lie algebra over \( k \) is the Lie subalgebra \( s_n(k) \) of scalar matrices. This corresponds to the first part of Exercise 3 on p10 of [13].

Similarly, the center \( Z(sl_n(k)) \) of \( sl_n(k) \) as a Lie algebra over \( k \) is the intersection of \( s_n(k) \) with \( sl_n(k) \). This consists of matrices of the form \( tI \), where \( t \in k \) satisfies \( n \cdot t = 0 \). If \( n \cdot 1 \) has a multiplicative inverse in \( k \), then \( Z(sl_n(k)) = \{0\} \). If \( n \cdot 1 = 0 \) in \( k \), then \( Z(sl_n(k)) = s_n(k) \). This corresponds to the second part of Exercise 3 on p10 of [13].

If \( r \in \{1, \ldots, n\} \), then

\[
[a, e_{r,r}] = \sum_{h \neq r} a_{h,r} e_{h,r} - \sum_{m \neq r} a_{r,m} e_{r,m},
\]

(10.4.4)

by (10.4.2) with \( q = r \). Let \( d_n(k) \) be the Lie subalgebra of \( gl_n(k) \) consisting of diagonal matrices, as in the previous section. Remember that \( a \in gl_n(k) \) is in the normalizer of \( d_n(k) \) in \( gl_n(k) \) if \( [a, b] \in d_n(k) \) for every \( b \in d_n(k) \), as in Section 9.8. In this case, it is easy to see that \( a \in d_n(k) \), using (10.4.4). This corresponds to part of Exercise 7 on p10 of [13].

If \( a \in d_n(k) \), then

\[
[a, e_{q,r}] = (a_{q,q} - a_{r,r}) e_{q,r}
\]

(10.4.5)

for every \( q, r \in \{1, \ldots, n\} \) with \( q \neq r \), by (10.4.3). This also works when \( q = r \), in which the right side is equal to 0, by (10.4.4). This corresponds to Exercise 6 on p5 of [13].

### 10.5 Solvability and traces

Let \( k \) be a field of characteristic 0, and let \( V \) be a vector space over \( k \) of positive finite dimension \( n \). Remember that the space \( gl(V) \) of linear mappings from \( V \) into itself is a Lie algebra over \( k \), with respect to the usual commutator bracket. Let \( A \) be a Lie subalgebra of \( gl(V) \), and suppose that \( A \) is solvable as a Lie algebra over \( k \). If \( T \in A \) and \( R \in [A, A] \), then

\[
(10.5.1) \quad \text{tr}(T \circ R) = 0,
\]
as in Theorem 7.1 on p42 of [24]. Here \([A, A]\) is the derived algebra of \(A\), as in Section 9.2, as usual.

Suppose for the moment that \(k\) is algebraically closed. Lie’s theorem implies that there is a flag \(F = \{V_j\}_{j=0}^n\) in \(V\) such that \(T(V_j) \subseteq V_j\) for every \(T \in A\) and \(j = 0, 1, \ldots, n\), as in Section 9.13. If \(R \in [A, A]\), then \(R(V_j) \subseteq V_{j-1}\) for \(j = 1, \ldots, n\), as in Section 9.11. This implies (10.5.1), using a basis for \(V\) that is compatible with \(F\). This corresponds to Exercise 7 on p21 of [13].

Now let \(k\) be any field of characteristic 0. Of course, \(V\) is isomorphic to \(k^n\) as a vector space over \(k\), and we may as well take \(V = k^n\). We can reformulate (10.5.1) in terms of matrices, as follows. If \(A_0\) is a Lie subalgebra of \(gl_n(k)\) that is solvable as a Lie algebra over \(k\), then

\[
(10.5.2) \quad \text{tr}(T_0 R_0) = 0
\]

for every \(T_0 \in A_0\) and \(R_0 \in [A_0, A_0]\). More precisely, this uses matrix multiplication and the trace on the space \(M_n(k)\) of \(n \times n\) matrices with entries in \(k\).

Let \(k_1\) be an algebraically closed field that contains \(k\) as a subfield, so that \(gl_n(k)\) is contained in \(gl_n(k_1)\). Let \(A_1\) be the linear span of \(A_0\) in \(gl_n(k_1)\), as a vector space over \(k_1\). It is easy to see that \(A_1\) is a Lie subalgebra of \(gl_n(k_1)\), as a Lie algebra over \(k_1\), because \(A_0\) is a Lie subalgebra of \(gl_n(k)\). Similarly, one can check that \(A_1\) is solvable as a Lie algebra over \(k_1\), because \(A_0\) is solvable as a Lie algebra over \(k\). It follows that

\[
(10.5.3) \quad \text{tr}(T_1 R_1) = 0
\]

for every \(T_1 \in A_1\) and \(R_1 \in [A_1, A_1]\), by the earlier argument for algebraically closed fields. This uses matrix multiplication and the trace on \(M_n(k_1)\), which contains \(M_n(k)\). This implies (10.5.2), because \(A_0 \subseteq A_1\), and hence \([A_0, A_0]\) is contained in \([A_1, A_1]\).

Let \(k\) be a commutative ring with a multiplicative identity element, and remember that \([gl_2(k), gl_2(k)] = sl_2(k)\), as in Section 10.1. Suppose that \(1 + 1 = 0\) in \(k\), so that the identity matrix in \(gl_2(k)\) is in \(sl_2(k)\). Remember that \(sl_2(k)\) is nilpotent as a Lie algebra over \(k\) in this case, as in Section 10.2, so that \(gl_2(k)\) is solvable as a Lie algebra over \(k\). In this situation, (10.5.2) would say that every element of \(gl_2(k)\) has trace 0, which is of course not the case.

### 10.6 Diagonalizable linear mappings

Let \(k\) be a field, let \(V\) be a vector space over \(k\), and let \(T\) be a linear mapping from \(V\) into itself. If \(\lambda \in k\), then put

\[
(10.6.1) \quad E_\lambda(T) = \{v \in V : T(v) = \lambda(v)\},
\]

which is a linear subspace of \(V\). As usual, \(\lambda\) is said to be an eigenvalue of \(T\) when \(E_\lambda(T) \neq \{0\}\), in which case the elements of \(E_\lambda(T)\) are said to be eigenvectors
of \( T \) corresponding to \( \lambda \). If \( R \) is another linear mapping from \( V \) into itself that commutes with \( T \) and \( v \in E_\lambda(T) \), then
\[ (10.6.2) \quad T(R(v)) = R(T(v)) = \lambda R(v). \]
This means that \( R(v) \in E_\lambda(T) \), so that
\[ (10.6.3) \quad R(E_\lambda(T)) \subseteq E_\lambda(T). \]

Suppose that \( \lambda_1, \ldots, \lambda_l \) are finitely many distinct elements of \( k \), and let \( r \in \{1, \ldots, n\} \) be given. Consider the linear mapping
\[ (10.6.4) \quad \prod_{j \neq r}(T - \lambda_j I) \]
on \( V \), where more precisely \( I \) is the identity mapping on \( V \), and the product is the composition of \( T_j - \lambda_j I \) for \( j = 1, \ldots, n \) and \( j \neq r \). This linear mapping is equal to 0 on \( E_\lambda(T) \) when \( j \neq r \), and it is equal to
\[ (10.6.5) \quad \prod_{j \neq r}(\lambda_r - \lambda_j) \]
times the identity mapping on \( E_\lambda(T) \). As before, \( (10.6.5) \) is the product of \( \lambda_r - \lambda_j \) for \( j = 1, \ldots, n \) and \( j \neq r \), which is a nonzero element of \( k \). If \( v_j \in E_\lambda(T) \) for each \( j = 1, \ldots, l \), then
\[ (10.6.6) \quad \sum_{j=1}^l v_j = 0, \]
then it is easy to see that \( v_r = 0 \) for every \( r = 1, \ldots, l \), by applying (10.6.4) to the sum on the left.

Suppose from now on in this section that \( V \) has positive finite dimension, as a vector space over \( k \). Note that \( T \) can have only finitely many distinct eigenvalues, and that the sum of the dimensions of the nontrivial eigenspaces of \( T \) is less than or equal to the dimension of \( V \), by the remark at the end of the preceding paragraph. If every element of \( V \) can be expressed as a sum of eigenvectors of \( T \), then \( T \) is said to be diagonalizable on \( V \). This means that \( V \) is the direct sum of the nontrivial eigenspaces of \( T \), and hence that there is a basis for \( V \) consisting of eigenvectors for \( T \).

Suppose that \( T \) is diagonalizable on \( V \), with distinct eigenvalues \( \lambda_1, \ldots, \lambda_l \). If \( r \in \{1, \ldots, n\} \), then (10.6.4) maps \( V \) onto \( E_{\lambda_r}(T) \). Let \( W \) be a linear subspace of \( V \) such that
\[ (10.6.7) \quad T(W) \subseteq W. \]
Observe that (10.6.4) maps \( W \) into itself for each \( r = 1, \ldots, n \). Of course, every \( w \in W \) can be expressed as a sum of eigenvectors of \( T \), by hypothesis. In this situation, these eigenvectors of \( T \) are also elements of \( W \), because (10.6.4) maps \( W \) into itself for each \( r \). This implies that the restriction of \( T \) to \( W \) is diagonalizable.
Let $R$ be another linear mapping from $V$ into itself that commutes with $T$ again. Thus $R$ maps $E_j(T)$ into itself for each $j = 1, \ldots, n$, as in (10.6.3). If $R$ is diagonalizable on $V$, then the restriction of $R$ to $E_j(T)$ is diagonalizable for every $r = 1, \ldots, n$, as in the previous paragraph. This implies that $R$ and $T$ are simultaneously diagonalizable on $V$, which is to say that there is a basis of $V$ consisting of vectors that are eigenvectors for both $R$ and $T$. In particular, it follows that $R + T$ and $R \circ T$ are diagonalized by the same basis for $V$.

10.7 Nilpotent vectors

Let $k$ be a field, and let $A$ be an associative algebra over $k$ with a multiplicative identity element $e$, where multiplication of $a, b \in A$ is expressed as $ab$. If $a \in A$ is nilpotent, $\lambda \in k$, and $\lambda \neq 0$, then $\lambda e + a$ has a multiplicative inverse in $A$. This follows from a remark in Section 3.1 when $\lambda = 1$, and otherwise one can reduce to that case.

Let $V$ be a vector space over $k$, and let $T$ be a linear mapping from $V$ into itself. Remember that

\begin{equation}
(I - T) \sum_{j=0}^{n} T^j = \left( \sum_{j=0}^{n} T^j \right) (I - T) = I - T^{n+1}
\end{equation}

for every nonnegative integer $n$, as in Section 3.1, where $I$ is the identity mapping on $V$. Suppose for the moment that for each $v \in V$ we have that $T^l(v) = 0$ for some nonnegative integer $l$, which implies that $T^j(v) = 0$ when $j \geq l$. This permits us to define

\begin{equation}
\sum_{j=0}^{\infty} T^j(v)
\end{equation}

as an element of $V$ for every $v \in V$, so that $\sum_{j=0}^{\infty} T^j$ is defined as a linear mapping from $V$ into itself. One can check that

\begin{equation}
(I - T) \sum_{j=0}^{\infty} T^j = \left( \sum_{j=0}^{\infty} T^j \right) (I - T) = I,
\end{equation}

using (10.7.1). This implies that $I - T$ is invertible on $V$, with

\begin{equation}
(I - T)^{-1} = \sum_{j=0}^{\infty} T^j.
\end{equation}

If $\lambda \in k$ and $\lambda \neq 0$, then $\lambda I - T = \lambda (I - (1/\lambda) T)$ is invertible on $V$ too, because $(1/\lambda) T$ satisfies the same condition on $V$.

Let $T$ be any linear mapping from $V$ into itself again, and put

\begin{equation}
\mathcal{E}_0(T) = \{ v \in V : T^l(v) = 0 \text{ for some } l \in \mathbb{Z}_+ \}.
\end{equation}

This is a linear subspace of $V$ that contains the kernel of $T$. If $v \in \mathcal{E}_0(T)$ and $v \neq 0$, and if $j$ is the smallest positive integer such that $T^j(v) = 0$, then $T^{j-1}(v)$
10.8. JORDAN–CHEVALLEY DECOMPOSITIONS

is a nonzero element of the kernel of $T$. Thus $\mathcal{E}_0(T) \neq \{0\}$ if and only if the kernel of $T$ is nontrivial. Note that $T$ maps $\mathcal{E}_0(T)$ into itself. If $\lambda \in k$ and $\lambda \neq 0$, then $\lambda I - T$ is invertible as a linear mapping on $\mathcal{E}_0(T)$, by the remarks in the previous paragraph. If $R$ is a linear mapping from $V$ into itself that commutes with $T$ and $v \in \mathcal{E}_0(T)$, then

$$T^l(R(v)) = R(T^l(v)) = 0$$

for some $l \in \mathbb{Z}_+$. This means that $R(v) \in \mathcal{E}_0(T)$, and hence

$$R(\mathcal{E}_0(T)) \subseteq \mathcal{E}_0(T).$$

Let $\lambda \in k$ be given, and put

$$\mathcal{E}_\lambda(T) = \mathcal{E}_0(T - \lambda I) = \{v \in V : (T - \lambda I)^l(v) = 0 \text{ for some } l \in \mathbb{Z}_+\}.$$

This is a linear subspace of $V$, which reduces to (10.7.5) when $\lambda = 0$. Of course, $E_\lambda(T) \subseteq \mathcal{E}_\lambda(T),

(10.7.9)

$$R(\mathcal{E}_\lambda(T)) \subseteq \mathcal{E}_\lambda(T),$$

as in (10.7.7). In particular, $T$ maps $\mathcal{E}_\lambda(T)$ into itself. If $\mu \in k$ and $\mu \neq \lambda$, then

$$\mu I - T = (\mu - \lambda) I - (T - \lambda I)$$

is invertible as a linear mapping from $\mathcal{E}_\lambda(T)$ into itself, by the analogous statement in the preceding paragraph.

Let $\lambda_1, \ldots, \lambda_m$ be finitely many distinct eigenvalues of $T$, and let $v_j \in \mathcal{E}_{\lambda_j}(T)$ be given for $j = 1, \ldots, m$. If

$$\sum_{j=1}^m v_j = 0,$$

then $v_r = 0$ for each $r = 1, \ldots, m$. To see this, one can apply suitable powers of $T - \lambda_j I$ to the sum on the left side of (10.7.12) for $j \neq r$, to get a product of powers of $T - \lambda_j I$ with $j \neq r$ applied to $v_r$. This can only be 0 when $v_r = 0$, because of the invertibility of $T - \lambda_j I$ on $\mathcal{E}_{\lambda_r}(T)$ when $j \neq r$.

10.8 Jordan–Chevalley decompositions

Let $k$ be a field, and let $V$ be a vector space over $k$ of positive finite dimension. If $T$ is a linear mapping from $V$ into itself, then one may wish to be able to express $T$ as

$$T = T_1 + T_2,$$

(10.8.1)
where $T_1$ is a diagonalizable linear mapping from $V$ into itself, $T_2$ is a nilpotent linear mapping from $V$ into itself, and $T_1$, $T_2$ commute with each other, and hence with $T$. Suppose that this is possible, so that $V$ corresponds to the direct sum of the eigenspaces of $T_1$, as in Section 10.6, and $T_2$ maps each of these eigenspaces into itself, because $T_1$ and $T_2$ commute. Let $\lambda$ be an eigenvalue of $T_1$, and let $E_\lambda(T_1)$ be the corresponding eigenspace of $T_1$, as before. On $E_\lambda(T_1)$, $T - \lambda I = T_2$, so that $T - \lambda I$ is nilpotent on $E_\lambda(T)$. This implies that

\[(10.8.2) \quad E_\lambda(T_1) \subset E_\lambda(T),\]

where the right side is as in (10.7.8). In particular, $\lambda$ is an eigenvalue of $T$ as well.

Let $\lambda_1, \ldots, \lambda_m$ be a list of all of the distinct eigenvalues of $T_1$, so that every element of $V$ can be expressed as a sum of elements of the corresponding eigenspaces $E_{\lambda_j}(T_1)$, $j = 1, \ldots, m$. An element of $V$ can be expressed in at most one way as a sum of elements of the $E_{\lambda_j}(T)$'s, $j = 1, \ldots, m$, as in the previous section. Using this and (10.8.2), we get that

\[(10.8.3) \quad E_{\lambda_j}(T_1) = E_{\lambda_j}(T)\]

for each $j = 1, \ldots, m$. One can also verify that $\lambda_1, \ldots, \lambda_m$ are all of the eigenvalues of $T$. If one already knows that every element of $V$ can be expressed as a sum of elements of the $E_{\lambda_j}(T)$'s, where $\lambda$ is an eigenvalue of $T$, then one can get $T_1$ and $T_2$ using the remarks in the previous section. It is well known that this holds when the characteristic polynomial of $T$ can be factored into a product of linear factors. In particular, this happens when $k$ is algebraically closed.

Note that $T_1$ and $T_2$ can be expressed as polynomials in $T$ with coefficients in $k$ and no constant term, as in part (b) of the proposition on p17 of [13], and Lemma 6.1 on p40 of [24]. Equivalently, this means that $T_1$ and $T_2$ can be expressed as linear combinations of positive powers of $T$. In particular, if $R$ is any linear mapping from $V$ into itself that commutes with $T$, then $R$ commutes with $T_1$ and $T_2$. Alternatively, if $R$ commutes with $T$, then one can check that $R$ commutes with $T_1$, using (10.7.10) and (10.8.3). This implies that $R$ commutes with $T_2$ as well, by (10.8.1).

Suppose that $T_3$ and $T_4$ are commuting linear mappings from $V$ into itself such that

\[(10.8.4) \quad T = T_3 + T_4,\]

$T_3$ is diagonalizable on $V$, and $T_4$ is nilpotent on $V$. Note that $T_3$ and $T_4$ commute with $T$, and hence with $T_1$ and $T_2$, as in the preceding paragraph. Of course,

\[(10.8.5) \quad T_1 - T_3 = T_4 - T_2,\]

by (10.8.1) and (10.8.4). In this situation, $T_1 - T_3$ is diagonalizable on $V$, as in Section 10.6. We also have that $T_4 - T_2$ is nilpotent on $V$, as in Section 9.7. The only linear mapping from $V$ into itself that is both diagonalizable and nilpotent is equal to 0 on $V$, so that $T_1 = T_3$ and $T_2 = T_4$. This is the uniqueness statement in [13, 24]. One could obtain $T_1 = T_3$ from (10.8.3) too.
10.9 SOME RELATED SITUATIONS

Suppose that \( W_0 \) and \( W \) are linear subspaces of \( V \) with

\[
T(W) \subseteq W_0 \subseteq W.
\]

Under these conditions,

\[
T_1(W), T_2(W) \subseteq W_0,
\]

as in part (c) of the proposition on p17 of [13], and Consequence 6.2 on p40 of [24]. This follows from the expressions for \( T_1 \) and \( T_2 \) in terms of polynomials in \( T \) mentioned earlier.

10.9 Some related situations

Let \( k \) be a field, and let \( V \) be a vector space over \( k \) of positive finite dimension. If \( A \) and \( B \) are linear mappings from \( V \) into itself, then put

\[
\text{ad}_A(B) = [A, B] = A \circ B - B \circ A,
\]

as before. This defines \( \text{ad}_A \) as a linear mapping from \( L(V) \) into itself, where \( L(V) \) is the space of linear mappings from \( V \) into itself. If \( A \) is diagonalizable on \( V \), then \( \text{ad}_A \) is diagonalizable on \( L(V) \). This can be seen using a basis for \( V \) consisting of eigenvectors for \( A \), to reduce to the analogous statement for matrices, as in Section 10.4. Similarly, if \( A \) is nilpotent on \( V \), then \( \text{ad}_A \) is nilpotent on \( L(V) \), as in Section 9.7. Let \( T, T_1, \) and \( T_2 \) be as in (10.8.1), so that

\[
\text{ad}_T = \text{ad}_{T_1} + \text{ad}_{T_2}.
\]

Note that \( \text{ad}_{T_1} \) is diagonalizable on \( L(V) \), and that \( \text{ad}_{T_2} \) is nilpotent on \( L(V) \), by the corresponding properties of \( T_1 \) and \( T_2 \) on \( V \), and the previous remarks. We also have that

\[
[\text{ad}_{T_1}, \text{ad}_{T_2}] = \text{ad}_{[T_1,T_2]} = 0,
\]

so that \( \text{ad}_{T_1} \) and \( \text{ad}_{T_2} \) commute on \( L(V) \), because \( T_1 \) and \( T_2 \) commute on \( V \). Thus (10.9.2) is the analogue of (10.8.1) for \( \text{ad}_T \), as in Lemma A on p18 of [13]. This also corresponds to Lemma 6.3 on p41 of [24], with \( p = q = 1 \).

Let \( A \) be an algebra over \( k \) in the strict sense, where multiplication of \( a, b \in A \) is expressed as \( ab \). Also let \( \delta \) be a derivation on \( A \). If \( a, b \in A \) are eigenvectors of \( \delta \) with eigenvalues \( \lambda, \mu \in k \), respectively, then

\[
\delta(ab) = \delta(a)b + a\delta(b) = (\lambda a)b + a(\mu b) = (\lambda + \mu)(ab),
\]

so that \( ab \) is an eigenvector of \( A \) with eigenvalue \( \lambda + \mu \) when \( ab \neq 0 \). It follows that the linear span in \( A \) of the eigenvectors of \( \delta \) forms a subalgebra of \( A \). This is stated in Exercise 12 on p6 of [13] for derivations on Lie algebras that come from the adjoint representation.

If \( \lambda \in k \), then put

\[
E_\lambda(\delta) = \{a \in A : (\delta - \lambda I)^{l}(a) = 0 \text{ for some } l \in \mathbb{Z}_+\},
\]
as in Section 10.7, where \( I \) is the identity mapping on \( A \). This is a linear subspace of \( A \) which is nontrivial exactly when \( \lambda \) is an eigenvalue of \( \delta \) on \( A \), as before. Let \( a, b \in A \) and \( \lambda, \mu \in k \) be given, and observe that

\[
(10.9.6) \quad (\delta - (\lambda + \mu) I)(a b) = \delta(a) b + a \delta(b) - \lambda a b - \mu a b = ((\delta - \lambda I)(a)) b + a ((\delta - \mu I)(b)).
\]

If \( a \in \mathcal{E}_\lambda(\delta) \) and \( b \in \mathcal{E}_\mu(\delta) \), then one can use this repeatedly to get that

\[
(10.9.7) \quad (\delta - (\lambda + \mu) I)^j(a b) = 0
\]

when \( j \) is sufficiently large, so that \( a b \in \mathcal{E}_{\lambda + \mu}(\delta) \). If \( a b \neq 0 \), then \( \mathcal{E}_{\lambda + \mu}(\delta) \neq \{0\} \), which implies that \( \lambda + \mu \) is an eigenvalue of \( \delta \), as before.

Suppose that \( k \) is algebraically closed, and that \( A \) has positive finite dimension as a vector space over \( k \). As in the previous section, there are linear mappings \( \delta_1, \delta_2 \) from \( A \) into itself such that

\[
(10.9.8) \quad \delta = \delta_1 + \delta_2,
\]

\( \delta_1 \) is diagonalizable on \( A \), \( \delta_2 \) is nilpotent on \( A \), and \( \delta_1, \delta_2 \) commute on \( A \). More precisely,

\[
(10.9.9) \quad E_\lambda(\delta_1) = \mathcal{E}_\lambda(\delta)
\]

for every \( \lambda \in k \), and in particular the eigenvalues of \( \delta \) and \( \delta_1 \) are the same. If \( a \in E_\lambda(\delta_1) = \mathcal{E}_\lambda(\delta) \) and \( b \in E_\mu(\delta_1) = \mathcal{E}_\mu(\delta) \) for some \( \lambda, \mu \in k \), then \( a b \) is an element of \( \mathcal{E}_{\lambda + \mu}(\delta) = E_{\lambda + \mu}(\delta_1) \), as in the preceding paragraph. Note that \( a b \) may be equal to 0, in which case \( \lambda + \mu \) need not be an eigenvalue of \( \delta, \delta_1 \). In this situation,

\[
(10.9.10) \quad \delta_1(a b) = (\lambda + \mu) a b = (\lambda a) b + a (\mu b) = \delta_1(a) b + a \delta_1(b).
\]

One can use this to get that \( \delta_1 \) is a derivation on \( A \), because \( A \) is spanned by the eigenspaces of \( \delta_1 \). It follows that \( \delta_2 \) is a derivation on \( A \) as well, by (10.9.8). This corresponds to Lemma B on p18-19 of [13].

### 10.10 Replicas

Let \( k \) be a field, let \( V \) be a vector space over \( k \) of positive finite dimension \( n \), and let \( A \) be a diagonalizable linear mapping from \( V \) into itself. Thus \( V \) corresponds to the direct sum of the nontrivial eigenspaces of \( A \). If \( \phi \) is a \( k \)-valued function on the set of eigenvalues of \( A \), then \( \phi(A) \) can be defined as a linear mapping from \( V \) into itself by putting \( \phi(A) = \lambda I \) on \( E_\lambda(A) \) for each eigenvalue \( \lambda \in k \) of \( A \). This corresponds to Definition 6.4 on p41 of [24], but with fewer conditions on \( \phi \) for the moment, as in a remark just after Definition 6.4 in [24]. Note that \( \phi(A) \) can be expressed as a polynomial in \( A \) with coefficients in \( k \), using a polynomial whose values at the eigenvalues of \( A \) are the same as the values of \( \phi \).
Equivalently, let $v_1, \ldots, v_n$ be a basis of $V$ consisting of eigenvectors of $A$ with eigenvalues $a_1, \ldots, a_n \in k$, so that

$$A(v_j) = a_j v_j \quad (10.10.1)$$

for every $j = 1, \ldots, n$. Using this basis, $\phi(A)$ can be characterized by

$$\phi(A)(v_j) = \phi(a_j) v_j \quad (10.10.2)$$

for each $j = 1, \ldots, n$. If $h, m \in \{1, \ldots, n\}$, then let $E_{h,m}$ be the linear mapping from $V$ into itself such that

$$E_{h,m}(v_j) = \delta_{m,j} v_h \quad (10.10.3)$$

for every $j = 1, \ldots, n$, where $\delta_{m,j}$ is as in Section 10.1. The $E_{h,m}$'s form a basis for the space $L(V)$ of linear mappings from $V$ into itself, as a vector space over $k$. Let $\text{ad}_A$ be the linear mapping from $L(V)$ into itself corresponding to $A$ as in (10.9.1). One can check that

$$\text{ad}_A(E_{h,m}) = (a_h - a_m) E_{h,m} \quad (10.10.4)$$

for every $h, m = 1, \ldots, n$, as in (10.4.5). Similarly,

$$\text{ad}_\phi(A)(E_{h,m}) = (\phi(a_h) - \phi(a_m)) E_{h,m} \quad (10.10.5)$$

for every $h, m = 1, \ldots, n$.

Suppose now that $\phi$ is defined on a subgroup of $k$, as a commutative group with respect to addition, that contains the eigenvalues of $A$. Suppose also that $\phi$ is a group homomorphism from this subgroup of $k$ into $k$, with respect to addition. Under these conditions, (10.10.5) implies that

$$\text{ad}_\phi(A)(E_{h,m}) = \phi(a_h - a_m) E_{h,m} \quad (10.10.6)$$

for every $h, m = 1, \ldots, n$. This means that

$$\text{ad}_\phi(A) = \phi(\text{ad}_A), \quad (10.10.7)$$

where the right side is obtained from $\text{ad}_A$ as a diagonalizable linear mapping from $L(V)$ into itself in the same way as before. This corresponds to Lemma 6.5 on p41 of [24], with $p = q = 1$, using a remark near the top of p41 in [24].

If 0 is not an eigenvalue of $A$, then $\phi(A)$ can be expressed as a polynomial in $A$ with coefficients in $k$ and constant term equal to 0, using a polynomial whose values at the eigenvalues of $A$ are the same as the values of $\phi$, and whose value at 0 is equal to 0. We can also do this when 0 is an eigenvalue of $A$, as long as $\phi(0) = 0$. Of course, $\phi(0) = 0$ automatically when $\phi$ is a homomorphism from an additive subgroup of $k$ into $k$, as in the preceding paragraph. In this case, we can apply the same argument to (10.10.7) on $L(V)$, to get that (10.10.7) can be expressed as a polynomial in $\text{ad}_A$ with coefficients in $k$ and constant term equal to 0.
If $k$ has characteristic 0, then $k$ may be considered as a vector space over the field $\mathbb{Q}$, the rational numbers. In this case, we may consider mappings from the linear span of the set of eigenvalues of $A$ in $k$, as a vector space over $\mathbb{Q}$, into $k$ that are linear over $\mathbb{Q}$, as in the proof of the lemma on p. 19 of [13]. In [24], one simply considers mappings $\phi$ from $k$ into itself that are linear over $\mathbb{Q}$. The corresponding linear mappings $\phi(A)$ are called replicas of $A$ by Chevalley, as mentioned on p. 42 of [24].

10.11 Two lemmas about traces

Let $k$ be an algebraically closed field of characteristic 0, and let $V$ be a vector space over $k$ of positive finite dimension $n$. Also let $T$ be a linear mapping from $V$ into itself, and let $T_1, T_2$ be the corresponding linear mappings from $V$ into itself discussed in Section 10.8, so that

\[(10.11.1)\quad T = T_1 + T_2,\]

$T_1$ is diagonalizable on $V$, $T_2$ is nilpotent on $V$, and $T_1, T_2$ commute. If $\phi$ is a mapping from the linear span of the eigenvalues of $T_1$ in $k$, as a vector space over $\mathbb{Q}$, into $k$ that is linear over $\mathbb{Q}$, then $\phi(T_1)$ can be defined as a linear mapping from $V$ into itself, as in the previous section. If

\[(10.11.2)\quad \text{tr}(T \circ \phi(T_1)) = 0\]

for all such $\phi$, then $T_1 = 0$, so that $T = T_2$ is nilpotent on $V$. This corresponds to Lemma 6.7 on p. 42 of [24], rephrased a bit in the way that it is used in the proof of the lemma on p. 19 of [13].

Remember that $T_2$ maps the eigenspaces of $T_1$ into themselves, because $T_1$ and $T_2$ commute. If $\phi$ is as in the preceding paragraph, then it follows that $T_2$ commutes with $\phi(T_1)$, by definition of $\phi(T_1)$. In particular, $T_2 \circ \phi(T_1)$ is nilpotent, and

\[(10.11.3)\quad \text{tr}(T_2 \circ \phi(T_1)) = 0.\]

Thus (10.11.2) is the same as saying that

\[(10.11.4)\quad \text{tr}(T_1 \circ \phi(T_1)) = 0.\]

Let $\lambda_1, \ldots, \lambda_l \in k$ be a list of the eigenvalues of $T_1$, and let $m_1, \ldots, m_l \in \mathbb{Z}_+$ be the dimensions of the corresponding eigenspaces of $T_1$ in $V$. Using (10.11.4) we get that

\[(10.11.5)\quad \sum_{j=1}^{l} m_j \cdot \lambda_j \cdot \phi(\lambda_j) = 0,\]

because $\phi(T_1) = \phi(\lambda_j) I$ on $E_{\lambda_j}(T_1)$ for each $j = 1, \ldots, l$, by construction. If we also ask that $\phi(\lambda_j)$ be in the subfield of $k$ corresponding to $\mathbb{Q}$ for each $j = 1, \ldots, l$, then we can apply $\phi$ to (10.11.5), to obtain that

\[(10.11.6)\quad \sum_{j=1}^{l} m_j \cdot \phi(\lambda_j)^2 = 0.\]
This means that \( \phi(\lambda_j) = 0 \) for each \( j = 1, \ldots, n \), because \( \phi(\lambda_j) \) is in the subfield of \( k \) corresponding to \( \mathbb{Q} \). This implies that 0 is the only eigenvalue of \( T_1 \), because the previous statement holds for all mappings \( \phi \) from the linear span of the eigenvalues of \( T_1 \) in \( k \) as a vector space over \( \mathbb{Q} \) into the subfield of \( k \) corresponding to \( \mathbb{Q} \) that are linear over \( \mathbb{Q} \). Thus \( T_1 = 0 \) on \( V \), because \( T_1 \) is diagonalizable on \( V \), by hypothesis, as desired.

If \( k \) is the field \( \mathbb{C} \) of complex numbers, then a variant of this argument was remarked by Bergman, as mentioned on p42 of [24]. Namely, if (10.11.5) holds with \( \phi \) equal to complex conjugation on \( \mathbb{C} \), then

\[
\sum_{j=1}^{l} m_j |\lambda_j|^2 = 0.
\]

(10.11.7)

This implies directly that 0 is the only eigenvalue of \( T_1 \), as desired.

Let \( A \) and \( B \) be linear subspaces of the space \( \mathcal{L}(V) \) of all linear mappings from \( V \) into itself, with

(10.11.8) \[ A \subseteq B. \]

Put

(10.11.9) \[ M = \{ T \in \mathcal{L}(V) : [T, B] \in A \text{ for every } B \in B \}. \]

If \( T \in M \) satisfies

(10.11.10) \[ \text{tr}(T \circ R) = 0 \]

for every \( R \in M \), then \( T \) is nilpotent on \( V \). This is the lemma stated on p19 of [13].

Let \( T \in M \) be given, and let \( T_1, T_2 \in \mathcal{L}(V) \) be as in (10.11.1) again. Also let \( \phi \) be a mapping from the linear span of the eigenvalues of \( T_1 \) in \( k \), as a vector space over \( \mathbb{Q} \), into \( k \) that is linear over \( \mathbb{Q} \), as before. Thus \( \phi(T_1) \) can be defined as a linear mapping from \( V \) into itself as in the previous section, and we would like to check that

(10.11.11) \[ \phi(T_1) \in M. \]

This means that (10.11.2) follows from (10.11.10), so that we can reduce to the previous statement.

If \( A \) is a linear mapping from \( V \) into itself, then \( \text{ad}_A \) denotes the corresponding linear mapping from \( \mathcal{L}(V) \) into itself, as in (10.9.1). The condition that \( T \in M \) means exactly that

(10.11.12) \[ \text{ad}_T(B) \subseteq A. \]

Remember that

(10.11.13) \[ \text{ad}_T = \text{ad}_{T_1} + \text{ad}_{T_2} \]

is the analogue of (10.11.1) for \( \text{ad}_T \) as a linear mapping from \( V \) into itself, as in Section 10.9. This implies that \( \text{ad}_{T_1} \) can be expressed as a polynomial in \( \text{ad}_T \) with coefficients in \( k \) and constant term equal to 0, as in Section 10.8. It follows from this and (10.11.12) that

(10.11.14) \[ \text{ad}_{T_1}(B) \subseteq A, \]
so that \( T_1 \in \mathcal{M} \).

We also have that
\[
\text{ad}_{\phi(T_1)} = \phi(\text{ad}_{T_1}),
\]
as in (10.10.7). Remember that (10.11.15) can be expressed as a polynomial in \( \text{ad}_{T_1} \) with coefficients in \( k \) and constant term equal to 0, as in the previous section. This implies that
\[
\text{ad}_{\phi(T_1)}(B) \subseteq A,
\]
because of (10.11.14). Thus (10.11.11) holds, as desired.

### 10.12 Cartan’s criterion

Let \( k \) be a field of characteristic 0, and let \( V \) be a vector space over \( k \) of positive finite dimension \( n \). Also let \( A \) be a Lie subalgebra of \( gl(V) \), and suppose that
\[
\text{tr}(T \circ R) = 0
\]
for every \( T \in A \) and \( R \in [A, A] \). Under these conditions, \( A \) is solvable as a Lie algebra over \( k \). This is part of Theorem 7.1 on p42 of [24]. Note that the converse was discussed in Section 10.5.

Suppose first that \( k \) is algebraically closed, which corresponds to the theorem on p20 in [13]. In order to show that \( A \) is solvable as a Lie algebra, it suffices to show that \([A, A]\) is nilpotent as a Lie algebra over \( k \). To do this, it is enough to show that every element of \([A, A]\) is nilpotent as a linear mapping on \( V \), as in the previous section.

Put
\[
\mathcal{M} = \{Z \in gl(V) : [Z, T] \in [A, A] \text{ for every } T \in A\},
\]
and observe that \( A \subseteq \mathcal{M} \). We would like to show that
\[
\text{tr}(R \circ Z) = 0
\]
for every \( R \in [A, A] \) and \( Z \in \mathcal{M} \). This will imply that every \( R \in [A, A] \) is nilpotent as a linear mapping on \( V \), as in the previous section.

In order to get (10.12.3), it is enough to check that
\[
\text{tr}([[T_1, T_2] \circ Z) = 0
\]
for every \( T_1, T_2 \in A \) and \( Z \in \mathcal{M} \). Remember that
\[
\text{tr}([[T_1, T_2] \circ Z) = -\text{tr}(T_2 \circ ([T_1, Z])),
\]
as in Section 7.8. If \( T_1 \in A \) and \( Z \in \mathcal{M} \), then \([T_1, Z] \in [A, A]\), by the definition (10.12.2) of \( \mathcal{M} \). If we also have that \( T_2 \in A \), then it follows that the right side of (10.12.5) is equal to 0, by (10.12.1). This implies (10.12.4), as desired.

Now let \( k \) be any field of characteristic 0. We may as well take \( V = k^n \), because \( V \) is isomorphic to \( k^n \) as a vector space over \( k \). The earlier statement
10.13. COMPARING RADICALS

can be reformulated in terms of matrices as saying that if $A_0$ is a Lie subalgebra of $gl_n(k)$ such that

$$\text{tr}(T_0 R_0) = 0$$

(10.12.6)

for every $T_0 \in A_0$ and $R_0 \in [A_0, A_0]$, then $A_0$ is solvable as a Lie algebra over $k$. This uses matrix multiplication and the trace on the space $M_n(k)$ of $n \times n$ matrices with entries in $k$.

Let $k_1$ be an algebraically closed field that contains $k$ as a subfield, so that $gl_n(k_1)$, as a vector space over $k_1$, is a Lie algebra over $k_1$, as mentioned in Section 10.5. Note that $[A_0, A_0]$ is an ideal in $A_0$ as a Lie algebra over $k$, and that $[A_1, A_1]$ is defined analogously as an ideal in $A_1$, as a Lie algebra over $k_1$, as in Section 9.2. One can check that $[A_1, A_1]$ is the same as the linear span of $[A_0, A_0]$ in $A_1$, as a vector space over $k_1$.

It follows that

$$\text{tr}(T_1 R_1) = 0$$

(10.12.7)

for every $T_1 \in A_1$ and $R_1 \in [A_1, A_1]$, because of (10.12.6). This uses matrix multiplication and the trace on $M_n(k_1)$. Thus $A_1$ is solvable as a Lie algebra over $k_1$, by the earlier argument for algebraically closed fields of characteristic 0. This implies that $A_0$ is solvable as a Lie algebra over $k$, as desired.

10.13 Comparing radicals

Let $k$ be a field, and let $(A, [\cdot, \cdot])$ be a Lie algebra over $k$. If $x \in A$, then $\text{ad}_x = \text{ad}_{A,x}$ is the linear mapping from $A$ into itself defined by

$$\text{ad}_x(z) = \text{ad}_{A,x}(z) = [x, z]$$

(10.13.1)

for every $z \in A$, as in Section 2.4. Suppose that $A$ has positive finite dimension as a vector space over $k$, so that the Killing form

$$\beta(x, y) = \text{tr}_A(\text{ad}_x \circ \text{ad}_y)$$

(10.13.2)

is defined as an element of $k$ for every $x, y \in A$. Remember that the radical

$$A^\beta = \{ x \in A : \beta(x, y) = 0 \text{ for every } y \in A \}$$

(10.13.3)

of (10.13.2) in $A$ is an ideal in $A$ as a Lie algebra over $k$, as in Section 7.11.

Of course, the space of linear mappings from $A$ into itself, as a vector space over $k$, is an associative algebra over $k$ with respect to composition of mappings, and hence a Lie algebra over $k$ with respect to the corresponding commutator bracket. Let $A_0$ be a Lie subalgebra of $A$. The image

$$\{ \text{ad}_x : x \in A_0 \}$$

(10.13.4)

of $A_0$ under the adjoint representation of $A$ is a Lie subalgebra of the Lie algebra of all linear mappings from $A$ into itself. Suppose that $\beta(x, y) = 0$ for every
x ∈ [A₀, A₀] and y ∈ A₀. If k has characteristic 0, then Cartan’s criterion implies that (10.13.4) is solvable as a Lie algebra over k. It follows that A₀ is solvable as a Lie algebra over k, because the kernel of the adjoint representation on A is the center of A. This corresponds to the corollary on p20 of [13] when A₀ = A.

The radical (10.13.3) of β in A automatically satisfies the conditions on A₀ mentioned in the preceding paragraph. If k has characteristic 0, then we get that Aβ is solvable as a Lie algebra over k. Let Rad A be the solvable radical of A, as in Section 9.4. It follows that

\[ A^β \subseteq \text{Rad} A \]  

when k has characteristic 0, because Aβ is an ideal in A as a Lie algebra.

Remember that A is said to be semisimple as a Lie algebra when Rad A = {0}. If A is semisimple and k has characteristic 0, then Aβ = {0}, by (10.13.5), which means that the Killing form (10.13.2) on A is nondegenerate. This corresponds to parts of the theorem on p22 of [13] and Theorem 2.1 on p44 of [24].

If B is an ideal in A and B is commutative as a Lie algebra over k, then B is contained in the radical (10.13.3) of β in A, as in Section 7.11. If the Killing form on A is nondegenerate, then it follows that B = {0}. Remember that A is semisimple as a Lie algebra when {0} is the only ideal in A that is commutative as a Lie algebra, as in Section 9.4. This means that A is semisimple as a Lie algebra when the Killing form on A is nondegenerate. This corresponds to the other parts of the theorem on p22 of [13] and Theorem 2.1 on p44 of [24].

10.14 Complementary ideals

Let k be a commutative ring with a multiplicative identity element, and let \((A, [,])\) be a Lie algebra over k. Suppose that A is isomorphic to kʳ as a module over k for some positive integer r, so that the Killing form \(\beta\) can be defined on A as in (10.13.2). Remember that \(\beta\) has the associativity or invariance property

\[ \beta([x, w], y) = \beta(x, [w, y]) \]  

for every \(w, x, y \in A\), as in Section 7.11. If B is a submodule of A, as a module over k, then put

\[ B^⊥ = B^⊥^β = \{ x \in A : \beta(x, y) = 0 \text{ for every } y \in B \}. \]  

This is a submodule of A, as a module over k, which is the same as the radical of \(\beta\) in A when \(B = A\). If B is an ideal in A as a Lie algebra over k, then it is easy to see that (10.14.2) is an ideal in A, using (10.14.1). This can also be obtained from statements in Sections 6.10 and 7.7.

Let us suppose from now on in this section that k is a field, and that A has positive finite dimension as a vector space over k. Let B be a linear subspace of A, so that (10.14.2) is a linear subspace of A too, as before. If the Killing form
\(\beta\) is nondegenerate as a bilinear form on \(A\), then the sum of the dimensions of \(B\) and \(B^\perp\) is equal to the dimension of \(A\), by standard arguments.

Now let \(B\) be an ideal in \(A\) as a Lie algebra over \(k\), so that (10.14.2) is an ideal in \(A\) too, as before. Thus

\[
A_0 = B \cap B^\perp
\]

is an ideal in \(A\) as well, and \(\beta(x, y) = 0\) for every \(x, y \in A_0\), by construction. If \(k\) has characteristic 0, then it follows that \(A_0\) is solvable as a Lie algebra over \(k\), as in the previous section. If \(A\) is semisimple as a Lie algebra over \(k\), then we get that \(A_0 = \{0\}\). This implies that

\[
[x, y] = 0
\]

for every \(x \in B\) and \(y \in B^\perp\), because \([x, y] \in B \cap B^\perp\).

If \(k\) has characteristic 0 and \(A\) is semisimple as a Lie algebra over \(k\), then the Killing form \(\beta\) is nondegenerate on \(A\), as in the previous section. Under these conditions, we have that

\[
B + B^\perp = \{x + y : x \in B, y \in B^\perp\} = A.
\]

More precisely, the dimension of \(B + B^\perp\) is equal to the sum of the dimensions of \(B\) and \(B^\perp\), because \(B \cap B^\perp = \{0\}\), as in the preceding paragraph. We have also seen that the sum of the dimensions of \(B\) and \(B^\perp\) is equal to the dimension of \(A\), because \(\beta\) is nondegenerate on \(A\). This implies (10.14.5), and hence that \(A\) is isomorphic to the direct sum of \(B\) and \(B^\perp\) as a Lie algebra over \(k\), because of (10.14.4).

This corresponds to the first step in the proof of the theorem on p23 in [13], and to Theorem 2.2 on p44 of [24]. Note that ideals in \(B\) and \(B^\perp\) are ideals in \(A\) in this situation, because of (10.14.4) and (10.14.5). This implies that \(B\) and \(B^\perp\) are semisimple as Lie algebras over \(k\), because \(A\) is semisimple. It follows that \(A/B\) is semisimple as a Lie algebra over \(k\), because it is isomorphic to \(B^\perp\).

### 10.15 Simple Lie algebras

Let \(k\) be a field. A Lie algebra \((A, [\cdot, \cdot])\) over \(k\) is said to be simple if \(A\) is not commutative as a Lie algebra, and if the only ideals in \(A\) are \(A\) itself and \(\{0\}\). See p6 of [13], and Definition 2.3 on p44 of [24]. Note that the trivial Lie algebra \(\{0\}\) is not considered to be simple, nor is the one-dimensional Lie algebra \(k\) with respect to the trivial Lie bracket. If \(A\) is a simple Lie algebra over \(k\), then the center \(Z(A)\) of \(A\) as a Lie algebra is trivial, because \(Z(A)\) is an ideal in \(A\), and \(A\) is not commutative as a Lie algebra. Similarly, if \(A\) is simple and \([A, A]\) is as in Section 9.2, then \([A, A] = A\), because \([A, A]\) is an ideal in \(A\), and \(A\) is not commutative. In particular, this means that \(A\) is not solvable as a Lie algebra, and in fact that \(A\) is semisimple as a Lie algebra.

Let \((A, [\cdot, \cdot])\) be a Lie algebra over \(A\) again, and let \(A_1, \ldots, A_n\) be finitely many ideals in \(A\). Suppose that every element of \(A\) can be expressed in a unique
way as the sum of elements of $A_1, \ldots, A_n$, so that $A$ corresponds to the direct sum of $A_1, \ldots, A_n$, as a vector space over $k$. Let $j, l \in \{1, \ldots, n\}$ be given, with $j \neq l$, so that $A_j \cap A_l = \{0\}$. If $a_j \in A_j$ and $a_l \in A_l$, then $[a_j, a_l] \in A_j \cap A_l$, because $A_j$ and $A_l$ are ideals in $A$. This implies that

\[(10.15.1) \quad [a_j, a_l] = 0,\]

because $A_j \cap A_l = \{0\}$. This means that $A$ corresponds to the direct sum of $A_1, \ldots, A_n$ as a Lie algebra over $k$, as remarked on p22 of [13]. Note that ideals in any $A_j$ as a Lie algebra over $k$ are ideals in $A$, because of (10.15.1).

Suppose from now on in this section that $A$ has characteristic 0, and that $A$ is a semisimple Lie algebra over $k$ with positive finite dimension as a vector space over $k$. Under these conditions, there are finitely many ideals $A_1, \ldots, A_n$ in $A$ such that $A$ corresponds to the direct sum of the $A_j$'s, as in the preceding paragraph, and $A_j$ is simple as a Lie algebra over $k$ for each $j = 1, \ldots, n$. Of course, if $A$ is already simple as a Lie algebra over $k$, then this holds with $n = 1$. Otherwise, $A$ corresponds to the direct sum of two proper ideals, each of which is semisimple as a Lie algebra over $k$, as in the previous section. One can repeat the process as needed until all of the ideals are simple as Lie algebras over $k$. At each step, any ideal in any of the ideals already obtained in $A$ is an ideal in $A$ too, as in the previous paragraph. This corresponds to the first part of the first theorem on p23 of [13], and to Corollary 1 on p45 of [24].

Remember that $[A_j, A_j] = A_j$ for each $j = 1, \ldots, n$, because the $A_j$’s are simple Lie algebras. This implies that $[A, A] = A$, as in the corollary on p32 in [13], and Corollary 2 on p45 of [24].

Let $C$ be a simple Lie algebra over $k$, and suppose that $\phi$ is a Lie algebra homomorphism from $A$ onto $C$. Note that $\phi(A_j)$ is an ideal in $C$ for each $j = 1, \ldots, n$, because $A_j$ is an ideal in $A$ and $\phi(A) = C$. It follows that for each $j = 1, \ldots, n$, $\phi(A_j)$ is either trivial or equal to $C$, because $C$ is simple. Clearly $\phi(A_{j_1}) \neq \{0\}$ for some $j_1 \in \{1, \ldots, n\}$, because $\phi(A) = C$. This can happen for at most one element of $\{1, \ldots, n\}$, by (10.15.1) and the noncommutativity of $C$. The restriction of $\phi$ to $A_{j_1}$ is injective, because $A_{j_1}$ is simple as a Lie algebra over $k$. This is the uniqueness property discussed on p45 of [24].

Let $B_0$ be a simple ideal in $A$. Note that the center $Z(A)$ of $A$ is trivial, because $A$ is semisimple. This implies that $[A, B_0] \neq \{0\}$, because $B_0 \neq \{0\}$. It follows that $[A_{j_0}, B_0] \neq \{0\}$ for some $j_0 \in \{1, \ldots, n\}$. Remember that $[A_{j_0}, B_0]$ is an ideal in $A$ that is contained in $A_{j_0}$ and $B_0$, because $A_{j_0}$ and $B_0$ are ideals in $A$, as in Section 9.2. Thus $A_{j_0} = [A_{j_0}, B_0] = B_0$, because $A_{j_0}$ and $B_0$ are simple. This is the uniqueness part of the first theorem on p23 of [13]. See also Theorem 4’ on p6 of [23].

If $B$ is a nontrivial proper ideal in $A$, then one can start the process for expressing $A$ as a direct sum of simple ideals using $B$, as before. Thus $B$ is also expressed as a direct sum of some of these ideals, as in the corollary on p23 of [13], and Corollary 3 on p45 of [24].
Chapter 11

Some examples and related properties

11.1 Simplicity and solvability

Let $k$ be a field, and remember that $x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, and $h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ form a basis for $sl_2(k)$, as a vector space over $k$. We also have that $[x, y] = h$, $[h, x] = 2 \cdot x$, and $[h, y] = -2 \cdot y$, as in Section 10.2. It is not difficult to check directly that $sl_2(k)$ is simple as a Lie algebra over $k$ when the characteristic of $k$ is different from 2, as on p6-7 of [13]. More precisely, suppose that $A$ is a nonzero ideal in $sl_2(k)$. Let

$$ax + by + ch$$

be a nonzero element of $A$, where $a, b, c \in k$. One can take the commutator of (11.1.1) with $x$ twice to get that $-2 \cdot b x \in A$, and similarly one can take the commutator of (11.1.1) with $y$ twice to get that $-2 \cdot a y \in A$. This implies that $x \in A$ when $b \neq 0$, and that $y \in A$ when $a \neq 0$, because the characteristic of $A$ is not 2. Otherwise, if $a = b = 0$, then $c \neq 0$, and it follows that $h \in A$. Thus $A$ contains at least one of $x$, $y$, or $h$. One can use this to get that $A$ contains each of $x$, $y$, and $h$, because the characteristic of $k$ is not 2. This means that $A = sl_2(k)$, as desired.

Let $k$ be any field, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. Suppose that the dimension of $A$ is 3, as a vector space over $k$, and that $[A, A] = A$, where $[A, A]$ is as in Section 9.2. Under these conditions, $A$ is simple as a Lie algebra over $k$, as in Exercise 5 on p10 of [13]. To see this, let $A_0$ be an ideal in $A$, so that the quotient $A/A_0$ is a Lie algebra over $k$ too. It is easy to see that

$$[A/A_0, A/A_0] = A/A_0$$

in this situation, as in Section 9.3. However, if $A/A_0$ has dimension 1 as a vector space over $k$, then the left side of (11.1.2) is \{0\}. Similarly, if $A/A_0$ has dimension 2 as a vector space over $k$, then the left side of (11.1.2) has dimension
less than or equal to 1. It follows that $A/A_0$ has dimension 0 or 3, which is to say that $A_0 = \{0\}$ or $A$, as desired.

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$ again. Suppose that the derived subalgebra $A^{(1)} = [A, A]$ is generated by two elements $a, b$ as a module over $k$, so that every element of $A^{(1)}$ can be expressed as a linear combination of $a$ and $b$, with coefficients in $k$. This implies that $A^{(2)} = [A^{(1)}, A^{(1)}]$ consists of multiples of $[a, b]_A$, with coefficients in $k$. It follows that $A^{(3)} = \{0\}$, so that $A$ is solvable as a Lie algebra. Of course, if $A^{(1)}$ consists of the multiples of a single element of $A$ with coefficients in $k$, then $A^{(2)} = \{0\}$.

Suppose that $k$ is a field, and that $A$ has dimension 3 as a vector space over $k$. If $[A, A]$ is a proper subset of $A$, then $[A, A]$ has dimension less than or equal to 2, as a vector space over $k$. This implies that $A$ is solvable as a Lie algebra, as in the preceding paragraph.

### 11.2 Traces on $M_n(k)$

Let $k$ be a commutative ring with a multiplicative identity element, and let $n$ be a positive integer. Remember that the space $M_n(k)$ of $n \times n$ matrices with entries in $k$ is an associative algebra over $k$ with respect to matrix multiplication, as in Section 2.8. If $a, x \in M_n(k)$, then

\[ L_a(x) = a x \]  

is defined as an element of $M_n(k)$ using matrix multiplication. This defines $L_a$ as a homomorphism from $M_n(k)$ into itself, as a module over $k$. If $b \in M_n(k)$ too, then

\[ L_a \circ L_b = L_{ab}, \]  

as in Section 2.8. It is not difficult to see that

\[ \text{tr}_{M_n(k)} L_a = n \cdot \text{tr} a \]  

for every $a \in M_n(k)$, where $\text{tr} a$ is the ordinary trace of $a$, as an $n \times n$ matrix with entries in $k$. The left side of (11.2.3) is the trace of $L_a$ as a module homomorphism from $M_n(k)$ into itself, as in Section 7.8. More precisely, the definition of $\text{tr}_{M_n(k)}$ uses the fact that $M_n(k)$ is isomorphic to $k^{n^2}$, as a module over $k$. It follows that

\[ \text{tr}_{M_n(k)} (L_a \circ L_b) = \text{tr}_{M_n(k)} L_{ab} = n \cdot \text{tr}(a b) \]  

for every $a, b \in M_n(k)$, using (11.2.2) in the first step.

Similarly, if $a \in M_n(k)$, then

\[ R_a(x) = x a \]  

defines $R_a$ as a module homomorphism from $M_n(k)$ into itself. If $b \in M_n(k)$ too, then

\[ R_a \circ R_b = R_{ab}, \]
as in Section 2.7. One can check that

\[(11.2.7) \quad \text{tr}_{M_n(k)} R_a = n \cdot \text{tr} a \]

for every \(a \in M_n(k)\). This implies that

\[(11.2.8) \quad \text{tr}_{M_n(k)}(R_a \circ R_b) = \text{tr}_{M_n(k)}(R_{ab}) = n \cdot \text{tr}(ba) = n \cdot \text{tr}(a b) \]

for every \(a, b \in M_n(k)\), using (11.2.6) in the first step.

Remember that \(L_a\) and \(R_b\) commute with each other on \(M_n(k)\) for every \(a, b \in M_n(k)\), as in Section 2.7. One can verify that

\[(11.2.9) \quad \text{tr}_{M_n(k)}(L_a \circ R_b) = (\text{tr} a)(\text{tr} b) \]

for every \(a, b \in M_n(k)\).

If \(a \in M_n(k)\), then

\[(11.2.10) \quad \text{ad}_a(x) = [a, x] = ax - xa = L_a(x) - R_a(x) \]

defines a module homomorphism from \(M_n(k)\) into itself, as usual. Equivalently,

\[(11.2.11) \quad \text{ad}_a = L_a - R_a. \]

If \(b \in M_n(k)\) too, then

\[(11.2.12) \quad \text{ad}_a \circ \text{ad}_b = (L_a - R_a) \circ (L_b - R_b) = L_a \circ L_b - L_a \circ R_b - R_a \circ L_b + R_a \circ R_b = L_{ab} - L_a \circ R_b - R_a \circ L_b + R_a \circ R_b. \]

It follows that

\[(11.2.13) \quad \text{tr}_{M_n(k)}(\text{ad}_a \circ \text{ad}_b) = \text{tr}_{M_n(k)}(L_a \circ R_b) - \text{tr}_{M_n(k)}(L_b \circ R_a) + \text{tr}_{M_n(k)}(R_{ab}) = n \cdot \text{tr}(ab) - 2 \cdot (\text{tr} a)(\text{tr} b) + n \cdot \text{tr}(ba) = 2n \cdot \text{tr}(ab) - 2 \cdot (\text{tr} a)(\text{tr} b). \]

Note that the right side is automatically equal to 0 when either \(a\) or \(b\) is a multiple of the identity matrix.

Of course, \(gl_n(k)\) is the same as \(M_n(k)\) as a module over \(k\), so that traces over \(gl_n(k)\) are the same as traces over \(M_n(k)\). If \(a \in sl_n(k)\), then let \(\text{ad}_{sl_n(k),a}\) be the mapping from \(sl_n(k)\) into itself defined in the usual way, which is the same as the restriction of \(\text{ad}_a\) to \(sl_n(k)\). Remember that \(sl_n(k)\) is isomorphic to \(k^{n^2-1}\) as a module over \(k\), and that \(gl_n(k)\) is isomorphic to the direct sum of \(sl_n(k)\) and \(k\), as modules over \(k\), as in Section 10.3. If \(a, b \in sl_n(k)\), then we have that

\[(11.2.14) \quad \text{tr}_{sl_n(k)}(\text{ad}_{sl_n(k),a} \circ \text{ad}_{sl_n(k),b}) = \text{tr}_{M_n(k)}(\text{ad}_a \circ \text{ad}_b), \]

as in Section 7.10. This implies that

\[(11.2.15) \quad \text{tr}_{sl_n(k)}(\text{ad}_{sl_n(k),a} \circ \text{ad}_{sl_n(k),b}) = 2n \cdot \text{tr}(ab), \]

by (11.2.13).
11.3 A nice criterion

Let \( k \) be a commutative ring with a multiplicative identity element, and let \((A, [\cdot, \cdot], A)\) be a Lie algebra over \( k \). Of course, the center \( Z(A) \) of \( A \) as a Lie algebra is a solvable ideal in \( A \). If every solvable ideal in \( A \) is contained in \( Z(A) \), then \( A \) is said to be reductive as a Lie algebra over \( k \), as in Exercise 5 on p30 and p102 of [13]. This means that the solvable radical \( \text{Rad} A \) of \( A \) is contained in \( Z(A) \), as in Section 9.4. In particular, this holds when \( A \) is commutative or semisimple as a Lie algebra.

Now let \( k \) be an algebraically closed field of characteristic 0, and let \( V \) be a vector space over \( k \) of positive dimension. Also let \( A \) be a Lie subsalgebra of the Lie algebra \( \text{gl}(V) \) of linear mappings from \( V \) into itself. Thus \( V \) may be considered as a module over \( A \), as a Lie algebra over \( k \). Suppose that \( V \) is irreducible as a module over \( A \). Under these conditions, \( A \) is reductive as a Lie algebra over \( k \), and every element of the center \( Z(A) \) of \( A \) is a scalar multiple of the identity mapping \( I \) on \( V \). In particular, the dimension of \( Z(A) \) as a vector space over \( k \) is less than or equal to 1. If \( A \subseteq \text{sl}(V) \), then \( A \) is semisimple as a Lie algebra.

Remember that the solvable radical \( \text{Rad} A \) of \( A \) is the maximal solvable ideal in \( A \), as in Section 9.4. Lie's theorem implies that there is a \( v_0 \neq 0 \) such that \( v_0 \) is an eigenvector of every element of \( \text{Rad} A \), as in Section 9.13. This means that there is a linear functional \( \lambda \) on \( \text{Rad} A \) such that

\[
T(v_0) = \lambda(T) v_0
\]

for every \( T \in \text{Rad} A \).

If \( R \in A \) and \( T \in \text{Rad} A \), then \([R, T] = R \circ T - T \circ R \in \text{Rad} A \), because \( \text{Rad} A \) is an ideal in \( A \). Under these conditions,

\[
\lambda([R, T]) = 0,
\]

as in Section 9.12.

Put

\[
V_\lambda = \{ v \in V : T(v) = \lambda(T) v \quad \text{for every } T \in \text{Rad} A \}.
\]

This is a linear subspace of \( V \) that contains \( v_0 \). Let \( R \in A \), \( T \in \text{Rad} A \), and \( v \in V_\lambda \) be given, and observe that

\[
T(R(v)) = R(T(v)) - ([R, T])(v) = R(\lambda(T) v) - \lambda([R, T]) v = \lambda(T) R(v).
\]

This uses the fact that \([R, T] \in \text{Rad} A \) in the second step, and (11.3.2) in the third step. It follows that \( R(v) \in V_\lambda \), so that \( R(V_\lambda) \subseteq V_\lambda \).

Because \( V \) is irreducible as a module over \( A \), we get that \( V_\lambda = V \). This means that every \( T \in \text{Rad} A \) is equal to \( \lambda(T) I \). In particular, \( \text{Rad} A \subseteq Z(A) \), which implies that \( \text{Rad} A = Z(A) \), as desired.

A basic property of reductive Lie algebras will be discussed in Section 13.8.
11.4 Ideals and structure constants

Let $k$ be a commutative ring with a multiplicative identity element, and let $n$ be a positive integer. Thus the space $k^n$ of $n$-tuples of elements of $k$ is a (free) module over $k$ with respect to coordinatewise addition and scalar multiplication. Let $c_{r,j;l}^r$ be an element of $k$ for every $j,l,r = 1,\ldots,n$. If $x,y \in k^n$, then let $[x,y]_{k^n}$ be the element of $k^n$ whose $r$th coordinate is given by

$$([x,y]_{k^n})_r = \sum_{j=1}^{n} \sum_{l=1}^{n} c_{r,j;l}^r x_j y_l$$

for every $r = 1,\ldots,n$. This defines a mapping from $k^n \times k^n$ into $k^n$ that is bilinear over $k$. Suppose that $c_{r,j;l}^r = -c_{r,l;j}^r$ and $c_{r,j;j}^r = 0$ for every $j,l,r = 1,\ldots,n$, so that $[x,x]_{k^n} = 0$ for every $x \in k^n$. Suppose also that the $c_{r,j;l}^r$’s satisfy (9.14.8), so that (11.4.1) satisfies the Jacobi identity, as before. This means that $k^n$ is a Lie algebra over $k$ with respect to (11.4.1).

Let $m < n$ be a positive integer, and let us use $k^m \{0\}$ to denote the space of $x \in k^n$ such that $x_j = 0$ when $j \geq m + 1$. Suppose that for every $r = 1,\ldots,n$,

$$c_{r,j;l}^r = 0 \quad \text{when} \quad j \geq m + 1 \text{ or } l \geq m + 1.$$

This implies that $k^m \{0\}$ is an ideal in $k^n$, as a Lie algebra over $k$ with respect to (11.4.1). Note that this condition is necessary for $k^m \{0\}$ to be an ideal in $k^n$ with respect to (11.4.1).

Let $K$ be a commutative associative algebra over $k$, and note that the space $K^n$ of $n$-tuples of elements of $K$ may be considered as a module over $k$ with respect to coordinatewise addition and scalar multiplication. If $a,a' \in K^n$, then let $[a,a']_{k^n}$ be the element of $K^n$ whose $r$th coordinate is equal to

$$([a,a']_{K^n})_r = \sum_{j=1}^{n} \sum_{l=1}^{n} c_{r,j;l}^r a_j a'_l$$

for each $r = 1,\ldots,n$. As in Section 9.14, $K^n$ is a Lie algebra over $k$ with respect to (11.4.4). We also have that $K^m \{0\}$ is an ideal in $K^n$ with respect to (11.4.4), because of (11.4.3). If $k^m \{0\}$ is solvable as a Lie algebra over $k$, then $K^m \{0\}$ is solvable as a Lie algebra over $k$ too, as before.

Suppose that $K$ has a multiplicative identity element $e$, so that $K^n$ may be considered as a module over $K$, and as a Lie algebra over $K$ with respect to (11.4.4). Remember that $t \mapsto te$ defines a ring homomorphism from $k$ into $K$, which leads to a homomorphism from $k^n$ into $K^n$, as Lie algebras over $k$. If $t \mapsto te$ is injective as a mapping from $k$ into $K$, then the corresponding mapping from $k^n$ into $K^n$ is injective.

Now let $k$ be a field, and let $(A,[\cdot,\cdot]_A)$ be a Lie algebra over $k$ of positive finite dimension $n$, as a vector space over $k$. Any choice of basis for $A$, as a vector
space over \( k \), leads to an isomorphism between \( A \) and \( k^n \), as vector spaces over \( k \). As in Section 9.14, there are structure constants \( c_{ji}^r \in k \) such that (11.4.1) corresponds to \( [\cdot,\cdot]_A \) with respect to this isomorphism.

Let \( K \) be a field that contains \( k \) as a subfield. Using the structure constants \( c_{ji}^r \) just mentioned, \( K^n \) becomes a Lie algebra over \( K \) with respect to (11.4.4). Any other choice of basis for \( A \) will lead to isomorphic Lie algebra structures on \( k^n \) and \( K^n \).

Suppose that \( A_0 \) is a proper nonzero ideal in \( A \), of dimension \( m \). We can choose a basis for \( A \) that contains a basis for \( A_0 \), so that \( A_0 \) corresponds to \( k^m \) with respect to this isomorphism. If \( A_0 \) is solvable as a Lie algebra over \( k \), then \( k^m \) is solvable as a Lie algebra over \( k \), and \( K^m \) is solvable as a Lie algebra over \( K \).

If \( K^n \) is semisimple as a Lie algebra over \( K \), then it follows that \( A \) is semisimple as a Lie algebra over \( k \). Similarly, if \( K^n \) is simple as a Lie algebra over \( K \), then \( A \) is simple as a Lie algebra over \( k \). This corresponds to part of Theorem 9 on p9 of [23] and the remark following it when \( k = \mathbb{R} \) and \( K = \mathbb{C} \).

### 11.5 Nondegeneracy and bilinear forms

Let \( k \) be a field, and let \( V \) be a vector space over \( k \) of finite positive dimension \( n \). If \( v_1, \ldots, v_n \) is a basis for \( V \) and \( l \in \{1, \ldots, n\} \), then there is a unique linear functional \( \lambda_l \) on \( V \) such that \( \lambda_l(v_j) \) is equal to 1 when \( j = l \) and to 0 when \( j \neq l \). If \( \mu \) is any linear functional on \( V \), then it is easy to see that

\[
\lambda_l = \sum_{l=1}^{n} \mu(v_l) \lambda_l.
\]

In fact, \( \lambda_1, \ldots, \lambda_n \) forms a basis for the dual space \( V' \) of all linear functionals on \( V \), as a vector space over \( k \).

Let \( \beta \) be a bilinear form on \( V \), so that

\[
\beta_w(v) = \beta(v, w)
\]

is a linear functional on \( V \) for each \( w \in W \). Observe that

\[
\beta_{vj} = \sum_{j=1}^{n} \beta_{vj}(v_j) \lambda_j = \sum_{j=1}^{n} \beta(v_j, v_l) \lambda_j
\]

for each \( l = 1, \ldots, n \), using (11.5.1) in the first step. Of course, \( w \mapsto \beta_w \) defines a linear mapping from \( V \) into \( V' \), and (11.5.3) expresses this linear mapping in terms of a matrix. In particular, \( \beta \) is nondegenerate as a bilinear form on \( V \) if and only if \( \beta(v_j, v_l) \) is invertible as an \( n \times n \) matrix with entries in \( k \).

Now let \( (A, [\cdot,\cdot]_A) \) be a Lie algebra over \( k \) with finite positive dimension \( n \) as a vector space over \( k \). Using a basis for \( A \), we get an isomorphism between \( A \) and \( k^n \), as vector spaces over \( k \). As before, there are structure constants \( c_{ji}^r \in k \) for \( j, l, r = 1, \ldots, n \) that satisfy (11.4.2) and (9.14.8) such that (11.4.1)
corresponds to $\mathbf{[\cdot,\cdot]}_{A}$ with respect to the isomorphism just mentioned. If $x \in k^n$, then
\begin{equation}
\text{ad}_{k^n,x}(z) = [x, z]_{k^n}
\end{equation}
defines a linear mapping from $k^n$ into itself, as in Section 2.4. The Killing form on $k^n$ is defined by
\begin{equation}
b_{k^n}(x, y) = \text{tr}_{k^n}(\text{ad}_{k^n,x} \circ \text{ad}_{k^n,y})
\end{equation}
for every $x, y \in k^n$, as in Section 7.9.

Let $K$ be a field that contains $k$ as a subfield, so that $K^n$ becomes a Lie algebra over $K$ with respect to (11.4.4). As before,
\begin{equation}
\text{ad}_{K^n,x} = [x, z]_{K^n}
\end{equation}
defines a linear mapping from $K^n$ into itself for each $x \in K^n$, and the Killing form on $K^n$ is defined by
\begin{equation}
b_{K^n}(x, y) = \text{tr}_{K^n}(\text{ad}_{K^n,x} \circ \text{ad}_{K^n,y})
\end{equation}
for every $x, y \in K^n$. If $x \in k^n$, then (11.5.4) and (11.5.6) are the same on $k^n$, because (11.4.1) and (11.4.4) are the same on $k^n$. This implies that (11.5.5) and (11.5.7) are the same when $x, y \in k^n$.

Let $u_1, \ldots, u_n$ be the standard basis elements of $k^n$, so that the $j$th coordinate of $u_l$ is equal to 1 when $j = l$ and to 0 when $j \neq l$. Note that $u_1, \ldots, u_n$ form a basis for $K^n$ as well, as a vector space over $K$. As in the preceding paragraph,
\begin{equation}
b_{k^n}(u_j, u_l) = b_{K^n}(u_j, u_l)
\end{equation}
for every $j, l = 1, \ldots, n$. It follows that (11.5.5) is nondegenerate as a bilinear form on $k^n$ if and only if (11.5.7) is nondegenerate as a bilinear form on $K^n$. More precisely, this happens exactly when the determinant of (11.5.8), as an $n \times n$ matrix with entries in $k$, is not 0. Remember that nondegeneracy of the Killing form is equivalent to semisimplicity of a finite-dimensional Lie algebra over a field of characteristic 0, as in Section 10.13. If $k$ has characteristic 0, then $K$ has characteristic 0, and we get that $k^n$ is semisimple as a Lie algebra over $k$ if and only if $K^n$ is semisimple as a Lie algebra over $K$. This corresponds to part of Theorem 9 on p9 of [23] when $k = \mathbb{R}$ and $K = \mathbb{C}$.

### 11.6 Bilinear forms and adjoints

Let $k$ be a commutative ring with a multiplicative identity element, and let $n$ be a positive integer. The space $k^n$ of $n$-tuples of elements of $k$ is a module over $k$ with respect to coordinatewise addition and scalar multiplication, as usual. Let $\beta(\cdot, \cdot)$ be a bilinear form on $k^n$, which is to say a mapping from $k^n \times k^n$ into $k$ that is bilinear over $k$. Remember that there is a unique $n \times n$ matrix $(\beta_{j,l})$ with entries in $k$ such that
\begin{equation}
\beta(x, y) = \sum_{j=1}^{n} \sum_{l=1}^{n} \beta_{j,l} x_j y_l
\end{equation}
for every \( x, y \in k^n \), as in Section 3.12. It is easy to see that \( \beta(\cdot, \cdot) \) is symmetric or antisymmetric as a bilinear form on \( k^n \) if and only if \( (\beta_{j,1}) \) is symmetric or antisymmetric as a matrix, respectively, as before. Similarly, \( \beta(x, x) = 0 \) for every \( x \in k^n \) if and only if \( (\beta_{j,j}) \) is antisymmetric and \( \beta_{j,j} = 0 \) for every \( j = 1, \ldots, n \). If \( 1 + 1 \) has a multiplicative inverse in \( k \), then this is equivalent to the antisymmetry of \( \beta(\cdot, \cdot) \) or \( (\beta_{j,j}) \).

Remember that \( M_n(k) \) is the space of \( n \times n \) matrices with entries in \( k \), which is an associative algebra over \( k \) with respect to matrix multiplication. If \( a \in M_n(k) \), then \( a^t \) denotes the transpose of \( a \), as usual. Let us also use \( \beta \) to denote \( (\beta_{j,1}) \) as an element of \( M_n(k) \), and let us suppose for the rest of the section that \( \beta \) is invertible in \( M_n(k) \). Put

\[
(11.6.2) \quad a^* = (\beta^{-1} a \beta)^t = \beta^t a^t (\beta^t)^{-1}
\]

defines a module homomorphism \( T_a \) from \( k^n \) into itself, as before. Remember that \( a \mapsto T_a \) defines an isomorphism from the space \( M_n(k) \) of \( n \times n \) matrices with entries in \( k \) onto the space of module homomorphisms from \( k^n \) into itself, as algebras over \( k \). By construction,

\[
(11.6.4) \quad \beta(T_a(x), y) = \beta(x, T_{a^*}(y))
\]

for every \( a \in M_n(k) \) and \( x, y \in k^n \), as in Section 3.12. More precisely, \( a^* \) is uniquely determined by this property, and we may call \( (T_a)^* = T_{a^*} \) the adjoint of \( T_a \) with respect to \( \beta \). This defines an opposite algebra automorphism on the space of module homomorphisms from \( k^n \) into itself.

Observe that

\[
(11.6.5) \quad (a^* t^*) = \beta^t (a^*)^t (\beta^t)^{-1} = \beta^t (a^t)^t (\beta^t)^{-1} = \beta^t (\beta^t)^{-1} (a^t) (\beta^t)^{-1}
\]

for every \( a \in M_n(k) \). If \( \beta \) is either symmetric or antisymmetric, then we get that

\[
(11.6.6) \quad (a^*)^* = a
\]

for every \( a \in M_n(k) \), and (11.6.2) is the same as

\[
(11.6.7) \quad a^* = \beta a^t \beta^{-1}.
\]

Thus \( a \mapsto a^* \) defines an algebra involution on \( M_n(k) \) in each of these two cases. Equivalently, \( T \mapsto T^* \) is an algebra involution on the space of module homomorphisms from \( k^n \) into itself in both cases. This may be considered as an instance of a remark in Section 3.14 as well.
11.7 Bilinear forms and symmetry conditions

Let us continue with the same notation and hypotheses as in the previous section, so that $\beta = (\beta_{j,i})$ is an invertible element of $M_n(k)$, and $\beta(\cdot, \cdot)$ is the corresponding bilinear form on $k^n$, as in (11.6.1). Remember that $a \in M_n(k)$ is invertible exactly when the determinant $\det a$ of $a$ is invertible in $k$, and that the determinant of the transpose of $a$ is the same as $\det a$. If $a$ is antisymmetric, then

$$\det a = \det a^t = \det(-a) = (-1)^n \det a.$$  \hfill (11.7.1)

If $a$ is invertible too, then it follows that $(-1)^n = 1$ in $k$. This implies that $-1 = 1$ in $k$ when $n$ is odd.

Let $a \in M_n(k)$ be given, and let $T_a$ be the corresponding module homomorphism from $k^n$ into itself, as in (11.6.3). Observe that

$$\beta(T_a(x), y)$$  \hfill (11.7.2)

defines another bilinear form on $k^n$. One can check that every bilinear form on $k^n$ corresponds to a unique $a \in M_n(k)$ in this way, because $\beta$ is supposed to be invertible. This defines an isomorphism between $M_n(k)$ and the space of bilinear forms on $k^n$, as modules over $k$. Equivalently, this defines an isomorphism between the space of module homomorphisms from $k^n$ into itself and the space of bilinear forms on $k^n$, as modules over $k$.

Remember that $T_a$ is said to be symmetric with respect to $\beta$ when

$$\beta(T_a(x), y) = \beta(x, T_a(y))$$  \hfill (11.7.3)

for every $x, y \in k^n$. This is equivalent to asking that $a$ be self-adjoint with respect to (11.6.2), which is to say that $a^* = a$. If $\beta(\cdot, \cdot)$ is symmetric as a bilinear form on $k^n$, then (11.7.3) holds if and only if

$$\beta(T_a(x), y) = \beta(T_a(y), x)$$  \hfill (11.7.4)

for every $x, y \in k^n$, which means that (11.7.2) is symmetric as a bilinear form on $k^n$. If $\beta(\cdot, \cdot)$ is antisymmetric as a bilinear form on $k^n$, then (11.7.3) holds if and only if

$$\beta(T_a(x), y) = -\beta(T_a(y), x)$$  \hfill (11.7.5)

for every $x, y \in k^n$, which means that (11.7.2) is antisymmetric as a bilinear form on $k^n$.

Similarly, $T_a$ is antisymmetric with respect to $\beta$ when

$$\beta(T_a(x), y) = -\beta(x, T_a(y))$$  \hfill (11.7.6)

for every $x, y \in k^n$. This is equivalent to asking that $a$ be anti-self-adjoint with respect to (11.6.2), so that $a^* = -a$. If $\beta(\cdot, \cdot)$ is symmetric as a bilinear form on $k^n$, then (11.7.6) holds if and only if (11.7.5) holds, so that (11.7.2) is antisymmetric as a bilinear form on $k^n$. If $\beta(\cdot, \cdot)$ is antisymmetric as a bilinear
form on \( k^n \), then (11.7.6) holds if and only if (11.7.4) holds, so that (11.7.2) is symmetric as a bilinear form on \( k^n \).

Of course, if \( \gamma = (\gamma_{j,t}) \in M_n(k) \), then

\[
(11.7.7) \quad \gamma(x, y) = \sum_{j=1}^n \sum_{t=1}^n \gamma_{j,t} x_j y_t
\]
defines a bilinear form on \( k^n \). As usual, \( \gamma(\cdot, \cdot) \) is symmetric or antisymmetric as a bilinear form on \( k^n \) exactly when \( \gamma \) is symmetric or antisymmetric as a matrix, respectively. Note that (11.7.2) corresponds to (11.7.7) with

\[
(11.7.8) \quad \gamma = \beta \alpha.
\]

Thus (11.7.2) is symmetric or antisymmetric as a bilinear form on \( k^n \) exactly when (11.7.8) is symmetric or antisymmetric as a matrix, respectively.

### 11.8 Traces and involutions

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( n \) be a positive integer. Also let \( a \mapsto a^* \) be an algebra involution on the algebra \( M_n(k) \) of \( n \times n \) matrices with entries in \( k \). Suppose that

\[
(11.8.1) \quad \text{tr} a^* = \text{tr} a
\]

for every \( a \in M_n(k) \) too, where \( \text{tr} a \in k \) is the usual trace of \( a \). Note that this condition holds automatically when \( a^* \) is as in (11.6.2). Remember that \( a \mapsto a^* \) is an involution on \( M_n(k) \) in that situation when \( \beta \) is symmetric or antisymmetric.

Put

\[
(11.8.2) \quad (a, b)_{M_n(k)} = \text{tr}(ab^*)
\]

for every \( a, b \in M_n(k) \), which defines a bilinear form on \( M_n(k) \). Observe that

\[
(11.8.3) \quad (a, b)_{M_n(k)} = \text{tr}(ab^*) = \text{tr}((b^*)^* a^*) = \text{tr}(ba^*) = (b, a)_{M_n(k)}
\]

for every \( a, b \in M_n(k) \), using (11.8.1) in the first step. If \( c \in M_n(k) \), then let

\[
(11.8.4) \quad L_c(a) = ca \quad \text{and} \quad R_c(a) = ac
\]

be the corresponding operators of left and right multiplication on \( M_n(k) \) by \( c \), respectively. Thus

\[
(11.8.5) \quad (L_c(a), b)_{M_n(k)} = \text{tr}(cabb^*) = \text{tr}(ab^* c) = \text{tr}(a(c^*b)^*) = (a, L_c^*(b))_{M_n(k)}
\]

for every \( a, b \in M_n(k) \). Similarly,

\[
(11.8.6) \quad (R_c(a), b)_{M_n(k)} = \text{tr}(a c b^*) = \text{tr}(a(b c^*)^*) = (a, R_c^*(b))_{M_n(k)}
\]
for every \(a, b \in M_n(k)\).

Let \(a, b\) be elements of \(M_n(k)\) again. If \(b\) is self-adjoint with respect to the given involution on \(M_n(k)\), so that \(b^* = b\), then

\[(11.8.7) \quad (a, b)_{M_n(k)} = \text{tr}(a b).\]

If \(a\) is self-adjoint and \(b\) is arbitrary, then

\[(11.8.8) \quad (a, b)_{M_n(k)} = (b, a)_{M_n(k)} = \text{tr}(b a) = \text{tr}(a b).\]

If \(a\) is arbitrary and \(b\) is anti-self-adjoint, so that \(b^* = -b\), then

\[(11.8.9) \quad (a, b)_{M_n(k)} = -\text{tr}(a b).\]

If \(c \in M_n(k)\), then put

\[(11.8.11) \quad C_c(a) = [c, a] = ca - ac = L_c(a) - R_c(a)\]

for every \(a \in M_n(k)\), so that \(C_c\) is a homomorphism from \(M_n(k)\) into itself, as a module over \(k\). Note that

\[(11.8.12) \quad (C_c, b)_{M_n(k)} = (a, C_c^*(b))_{M_n(k)}\]

for every \(a, b \in M_n(k)\), by (11.8.5) and (11.8.6). Of course,

\[(11.8.13) \quad (C_c(a))^* = ([c, a])^* = -[c^*, a^*] = -C_{c^*}(a^*)\]

for every \(a \in M_n(k)\).

Suppose that 1 + 1 has a multiplicative inverse in \(k\). If \(a \in M_n(k)\) is self-adjoint and \(b \in M_n(k)\) is anti-self-adjoint, then

\[(11.8.14) \quad (a, b)_{M_n(k)} = 0,\]

by (11.8.8) and (11.8.9). This also holds when \(a\) is anti-self-adjoint and \(b\) is self-adjoint, by (11.8.3).

### 11.9 Inner products over ordered fields

Let \(k\) be an ordered field, as in Section 8.12, and let \(V\) be a vector space over \(k\). Also let \(\langle v, w \rangle_V\) be an inner product on \(V\), which is to say a symmetric bilinear form on \(V\) that is positive definite, in the sense that

\[(11.9.1) \quad \langle v, v \rangle_V > 0\]

for every \(v \in V\) with \(v \neq 0\). As usual, \(v, w \in V\) are said to be orthogonal with respect to \(\langle \cdot, \cdot \rangle_V\) when

\[(11.9.2) \quad \langle v, w \rangle_V = 0.\]
Suppose that $e_1, \ldots, e_n$ are finitely many nonzero vectors in $V$ that are pairwise orthogonal, so that
\begin{equation}
\langle e_j, e_l \rangle_V = 0 \quad (11.9.3)
\end{equation}
when $j \neq l$. If $v \in V$, then put
\begin{equation}
w = \sum_{j=1}^{n} \frac{\langle v, e_j \rangle_V}{\langle e_j, e_j \rangle_V} e_j,
\end{equation}
which is an element of the linear span of $e_1, \ldots, e_n$. Observe that
\begin{equation}
\langle w, e_l \rangle_V = \langle v, e_l \rangle_V \quad (11.9.5)
\end{equation}
for every $l = 1, \ldots, n$. Equivalently, $\langle w - v, e_l \rangle_V = 0$ for each $l = 1, \ldots, n$, so that
\begin{equation}
\langle w - v, u \rangle_V = 0 \quad (11.9.6)
\end{equation}
for every $u$ in the linear span of $e_1, \ldots, e_n$ in $V$. If $v$ is in the linear span of $e_1, \ldots, e_n$ in $V$, then we get that
\begin{equation}
v = w,
\end{equation}
because $w$ is in the linear span of $e_1, \ldots, e_n$ in $V$, by construction.

Suppose from now on in this section that $V$ has positive finite dimension $n$, as a vector space over $k$. In this case, there are nonzero pairwise-orthogonal vectors $e_1, \ldots, e_n$ in $V$ whose linear span is equal to $V$, so that they form a basis for $V$ as a vector space over $k$. This can be obtained from the Gram–Schmidt process, using the remarks in the preceding paragraph. If $v \in V$, then we get that
\begin{equation}
v = \sum_{j=1}^{n} \frac{\langle v, e_j \rangle_V}{\langle e_j, e_j \rangle_V} e_j,
\end{equation}
as before. Of course, if $k = \mathbb{R}$, then we can take the $e_j$’s to be orthonormal in $V$, so that $\langle e_j, e_j \rangle_V = 1$ for every $j = 1, \ldots, n$.

Remember that $\mathcal{L}(V)$ denotes the algebra of linear mappings from $V$ into itself. If $T \in \mathcal{L}(V)$, then there is a unique adjoint mapping $T^* \in \mathcal{L}(V)$ such that
\begin{equation}
\langle T(v), w \rangle_V = \langle v, T^*(w) \rangle_V \quad (11.9.9)
\end{equation}
for every $v, w \in V$, as usual. We have also seen that $T \mapsto T^*$ is an algebra involution on $\mathcal{L}(V)$. Observe that
\begin{equation}
T(e_l) = \sum_{j=1}^{n} \frac{\langle T(e_l), e_j \rangle_V}{\langle e_j, e_j \rangle_V} e_j \quad (11.9.10)
\end{equation}
for each $l = 1, \ldots, n$, so that
\begin{equation}
\operatorname{tr}_V T = \sum_{j=1}^{n} \frac{\langle T(e_j), e_j \rangle_V}{\langle e_j, e_j \rangle_V} \quad (11.9.11)
\end{equation}
11.9. INNER PRODUCTS OVER ORDERED FIELDS

This implies that

\[
\text{tr}_V T^* = \sum_{j=1}^n \frac{\langle T^*(e_j), e_j \rangle_V}{\langle e_j, e_j \rangle_V} = \sum_{j=1}^n \frac{\langle e_j, T(e_j) \rangle_V}{\langle e_j, e_j \rangle_V} = \text{tr}_V T.
\]

(11.9.12)

If \( T_1, T_2 \in \mathcal{L}(V) \), then put

\[
\langle T_1, T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (T_1 \circ T_2^*) = \text{tr}_V (T_2^* \circ T_1),
\]

which defines a bilinear form on \( \mathcal{L}(V) \). Using (11.9.12), we get that

\[
\langle T_1, T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (T_1 \circ T_2^*)^* = \text{tr}_V (T_2 \circ T_1) = \langle T_2, T_1 \rangle_{\mathcal{L}(V)},
\]

(11.9.13)

as in the previous section. Observe that

\[
\langle T; T \rangle_{\mathcal{L}(V)} = \text{tr}_V (T \circ T) \geq \sum_{j=1}^n \frac{\langle T^*(T(e_j)), e_j \rangle_V}{\langle e_j, e_j \rangle_V} = \sum_{j=1}^n \frac{\langle T(e_j), T(e_j) \rangle_V}{\langle e_j, e_j \rangle_V}
\]

for every \( T \in \mathcal{L}(V) \). If \( T \neq 0 \), then it follows that

\[
\langle T, T \rangle_{\mathcal{L}(V)} > 0,
\]

(11.9.16)

because each of the terms on the right side of (11.9.15) is greater than or equal to 0 in \( k \), and at least one term is strictly positive. Thus (11.9.13) defines an inner product on \( \mathcal{L}(V) \), as a vector space over \( k \).

Suppose that \( T_1 \in \mathcal{L}(V) \) is self-adjoint and \( T_2 \in \mathcal{L}(V) \) is anti-self-adjoint, so that \( T_1^* = T_1 \) and \( T_2^* = -T_2 \). Under these conditions, we have that

\[
\langle T_1, T_2 \rangle_{\mathcal{L}(V)} = -\text{tr}_V (T_2 \circ T_1) = -\langle T_2, T_1 \rangle_{\mathcal{L}(V)}.
\]

(11.9.17)

This means that

\[
\langle T_1, T_2 \rangle_{\mathcal{L}(V)} = 0.
\]

(11.9.18)

If \( A \in \mathcal{L}(V) \), then put

\[
L_A(T) = A \circ T \quad \text{and} \quad R_A(T) = T \circ A
\]

for every \( T \in \mathcal{L}(V) \), as before. Observe that

\[
\langle L_A(T_1), T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (A \circ T_1 \circ T_2^*) = \text{tr}_V (T_1 \circ T_2^* \circ A) = \langle T_1, L_A(T_2) \rangle_{\mathcal{L}(V)}
\]

(11.9.19)

and

\[
\langle R_A(T_1), T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (T_1 \circ A \circ T_2^*) = \text{tr}_V (T_1 \circ (T_2 \circ A)^*) = \langle T_1, R_A(T_2) \rangle_{\mathcal{L}(V)}
\]

(11.9.20)
for every $T_1, T_2 \in \mathcal{L}(V)$. Put

$$C_A(T) = [A, T] = A \circ T - T \circ A = L_A(T) - R_A(T)$$

for every $T \in \mathcal{L}(V)$. Thus

$$(11.9.22) \quad \langle C_A(T_1), T_2 \rangle_{\mathcal{L}(V)} = \langle T_1, C_A(T_2) \rangle_{\mathcal{L}(V)}$$

for every $T_1, T_2 \in \mathcal{L}(V)$, by (11.9.21) and (11.9.22).

### 11.10 Self-adjoint Lie subalgebras

Let us continue with the same notation and hypotheses as in the previous section. As before, we use $gl(V)$ to denote $\mathcal{L}(V)$, considered as a Lie algebra over $k$ with respect to the corresponding commutator bracket. If $R \in gl(V)$, then put

$$C_R(T) = [R, T] = R \circ T - T \circ R$$

for every $T \in gl(V)$, as in (11.9.22), which is the same as $\text{ad}_{gl(V)}(T)$. Remember that

$$(11.10.1) \quad \langle C_R(T_1), T_2 \rangle_{\mathcal{L}(V)} = \langle T_1, C_R(T_2) \rangle_{\mathcal{L}(V)}$$

for every $T_1, T_2 \in gl(V)$, as in (11.9.23). This means that $C_R$ is the adjoint of $C_R$ as a linear mapping from $gl(V) = \mathcal{L}(V)$ into itself, with respect to the inner product (11.9.13).

Let $A$ be a Lie subalgebra of $gl(V)$, as a Lie algebra over $k$. If $R \in A$, then $C_R$ maps $A$ into itself, and the restriction of $C_R$ to $A$ is the same as $\text{ad}_{A,R}$. If $R_1, R_2 \in A$, then

$$(11.10.3) \quad \text{tr}_A(\text{ad}_{A,R_1} \circ \text{ad}_{A,R_2}) = \text{tr}_A(C_{R_1} \circ C_{R_2})$$

is the same as the Killing form on $A$ evaluated at $R_1, R_2$. This is the trace of $\text{ad}_{A,R_1} \circ \text{ad}_{A,R_2}$ as a linear mapping from $A$ into itself, where $A$ is considered as a finite-dimensional vector space over $k$. In the right side of (11.10.3), one should use the restrictions of $C_{R_1}$ and $C_{R_2}$ to $A$, as indicated by taking the trace is taken over $A$.

If $A$ is any subset of $\mathcal{L}(V)$, then put

$$(11.10.4) \quad A^* = \{ T^* : T \in A \},$$

which is a subset of $\mathcal{L}(V)$ as well. Let us say that $A$ is self-adjoint as a subset of $\mathcal{L}(V)$ when $A^* = A$. Let $A$ be a Lie subalgebra of $gl(V)$ again, and suppose that $A$ is self-adjoint as a subset of $\mathcal{L}(V)$. If $R \in A$, then $R^* \in A$, so that $C_R$ and $C_{R^*}$ both map $A$ into itself. The restriction of $C_{R^*}$ to $A$ is the same as the adjoint of the restriction of $C_R$ to $A$, with respect to the restriction of the inner product (11.9.13) on $\mathcal{L}(V)$ to $A$.

Let $\mathcal{L}(A)$ be the space of all linear mappings from $A$ into itself, as usual. The restriction of the inner product (11.9.13) on $\mathcal{L}(V)$ to $A$ defines an inner
11.11. COMPARING INVOLUTIONS

Let \( \mathcal{A} \) be a commutative group, where the group operations are expressed additively, and let
\[
(\mathcal{A}^{1}) \quad x \mapsto x^{*_{1}}
\]
and
\[
(\mathcal{A}^{2}) \quad x \mapsto x^{*_{2}}
\]
be group homomorphisms from \( \mathcal{A} \) into itself. Suppose for the moment that these two group homomorphisms commute on \( \mathcal{A} \), so that
\[
(\mathcal{A}^{3}) \quad (x^{*_{1}})^{*_{2}} = (x^{*_{2}})^{*_{1}}
\]
for every \( x \in \mathcal{A} \). Of course, one can define self-adjointness and anti-self-adjointness of elements of \( \mathcal{A} \) with respect to \( (\mathcal{A}^{1}) \) and \( (\mathcal{A}^{2}) \) in the usual way. If \( x \in \mathcal{A} \) is self-adjoint with respect to \( (\mathcal{A}^{1}) \), then it follows that \( x^{*_{1}} \) is self-adjoint with respect to \( (\mathcal{A}^{2}) \) as well. Similarly, if \( x \) is anti-self-adjoint with respect to \( (\mathcal{A}^{1}) \), then \( x^{*_{1}} \) is anti-self-adjoint with respect to \( (\mathcal{A}^{2}) \).

Now let \( k \) be a commutative ring with a multiplicative identity element, and let \( \mathcal{A} \) be an associative algebra over \( k \) with a multiplicative identity element \( e \), where multiplication of \( a, b \in \mathcal{A} \) is expressed as \( ab \). Suppose that \( (\mathcal{A}^{1}) \) is an algebra involution on \( \mathcal{A} \), and let \( c \) be an invertible element of \( \mathcal{A} \). Put
\[
(\mathcal{A}^{4}) \quad x^{*_{2}} = c^{-1} x^{*_{1}} c
\]
for every \( x \in \mathcal{A} \). This defines (11.11.2) as an opposite algebra automorphism on \( \mathcal{A} \). Let us suppose from now on in this section that

\[(11.11.5) \quad c^{*1} = c \]

or

\[(11.11.6) \quad c^{*1} = -c.\]

In either case, (11.11.2) is an algebra involution on \( \mathcal{A} \) as well, as in Section 3.14. Note that

\[(11.11.7) \quad c^{*2} = c \]

when (11.11.5) holds, and that

\[(11.11.8) \quad c^{*2} = -c \]

when (11.11.6) holds.

If \( x \in \mathcal{A} \), then

\[(x^{*1})^{*2} = c^{-1} (x^{*1})^{*1} c = c^{-1} x c \]

and

\[(11.11.9) \quad (x^{*2})^{*1} = c^{-1} x^{*1} c = c^{*1} (x^{*1})^{*1} (c^{*1})^{-1} = c^{*1} x (c^{*1})^{-1}.\]

This reduces to

\[(11.11.10) \quad (x^{*2})^{*1} = c x c^{-1} \]

when \( c \) satisfies (11.11.5) or (11.11.6). If

\[(11.11.11) \quad c^2 = t c \]

for some \( t \in k \) with a multiplicative inverse in \( k \), then we get that (11.11.3) holds. If \( k = \mathbb{C} \), then (11.11.1) may be conjugate-linear, as usual.

Let \( \mathcal{A}_1 \) be a subalgebra of \( \mathcal{A} \), and suppose that (11.11.1) maps \( \mathcal{A}_1 \) into itself. There are some situations where \( c \) is not necessarily in \( \mathcal{A}_1 \), but \( c^{*1} x c \in \mathcal{A}_1 \) for every \( x \in \mathcal{A}_1 \). If \( x \in \mathcal{A}_1 \), then it follows that (11.11.4) is an element of \( \mathcal{A}_1 \) as well.

### 11.12 Some projections

Let \( k \) be a field, let \( V \) be a vector space over \( k \), and let \( b(v, w) \) be a bilinear form on \( V \). Also let \( V_0 \) be a finite-dimensional linear subspace of \( V \), and suppose that the restriction of \( b(v, w) \) to \( v, w \in V_0 \) is nondegenerate on \( V_0 \). If \( z \in V \), then put

\[(11.12.1) \quad \lambda_z(v) = b(v, z) \]

for every \( v \in V_0 \), which defines a linear functional on \( V_0 \). Because \( b(\cdot, \cdot) \) is nondegenerate on \( V_0 \), there is a unique element \( P_0(z) \) of \( V_0 \) such that

\[(11.12.2) \quad \lambda_z(v) = b(v, P_0(z))\]
for every \( v \in V_0 \). Equivalently, this means that
\[
(11.12.3) \quad b(v, z - P_0(z)) = 0
\]
for every \( v \in V \). It is easy to see that \( P_0 \) defines a linear mapping from \( V \) into \( V_0 \). More precisely,
\[
(11.12.4) \quad P_0(z) = z
\]
when \( z \in V_0 \), so that \( P_0 \) maps \( V \) onto \( V_0 \). It follows that
\[
(11.12.5) \quad P_0 \circ P_0 = P_0
\]
on \( V \), so that \( P_0 \) is a projection on \( V \).

Let \( Z_0 \) be the kernel of \( P_0 \), which is a linear subspace of \( V \). Note that
\[
(11.12.6) \quad V_0 \cap Z_0 = \{0\},
\]
by (11.12.4). If \( z \in V \), then
\[
(11.12.7) \quad z - P_0(z) \in Z_0,
\]
by (11.12.5). Thus \( z \in V_0 + Z_0 \), so that \( V = V_0 + Z_0 \). This shows that \( V \) corresponds to the direct sum of \( V_0 \) and \( Z_0 \), as a vector space over \( k \). If \( z \in Z_0 \), then
\[
(11.12.8) \quad b(v, z) = 0
\]
for every \( v \in V_0 \), by (11.12.3). This is equivalent to (11.12.3), because of (11.12.7).

Suppose that \( b(\cdot, \cdot) \) is either symmetric or antisymmetric on \( V \). This implies that
\[
(11.12.9) \quad b(z, v) = 0
\]
for every \( v \in V_0 \) and \( z \in Z_0 \), by (11.12.8). If \( u, w \in V \), then \( P_0(u), P_0(w) \in V_0 \), \( u - P_0(u), w - P_0(w) \in Z_0 \), and hence
\[
(11.12.10) \quad b(P_0(u), w - P_0(w)) = b(u - P_0(u), P_0(w)) = 0,
\]
by (11.12.8) and (11.12.9). It follows that
\[
(11.12.11) \quad b(u, w) = b(P_0(u), P_0(w)) + b(u - P_0(u), w - P_0(w)).
\]
Thus \( b(\cdot, \cdot) \) corresponds, as a bilinear form on \( V \), to the bilinear form on the direct sum of \( V_0 \) and \( Z_0 \) obtained from the restrictions of \( b(\cdot, \cdot) \) to \( V_0 \) and \( Z_0 \).

Suppose for the moment that
\[
(11.12.12) \quad b(w, w) = 0
\]
for every \( v \in V \). Remember that this implies that \( b(\cdot, \cdot) \) is antisymmetric on \( V \), and that the converse holds when the characteristic of \( k \) is different from 2, as in Section 2.1. Let \( x, y \) be elements of \( V \) such that
\[
(11.12.13) \quad b(x, y) \neq 0.
\]
This implies that \( x \) and \( y \) are linearly independent in \( V \), because of (11.12.12). Thus the linear span \( V_0 \) of \( x \) and \( y \) in \( V \) is a two-dimensional linear subspace of \( V \). It is easy to see that the restriction of \( b(\cdot, \cdot) \) to \( V_0 \) is nondegenerate on \( V_0 \). It follows that there is a complementary linear subspace \( Z_0 \) of \( V_0 \) with the properties discussed in the previous paragraphs.

Suppose that \( V \) has finite dimension as a vector space over \( k \). If \( b(\cdot, \cdot) \neq 0 \) on \( V \), then there are \( x, y \in V \) as in the preceding paragraph. Repeating the process, we get that \( V \) corresponds to the direct sum of finitely many two-dimensional linear subspaces on which \( b(\cdot, \cdot) \) is nondegenerate, and possibly an additional linear subspace \( W \) such that (11.12.12) holds for every \( w \in W \). As before, \( b(\cdot, \cdot) \) corresponds to the bilinear form on the direct sum obtained from the restrictions of \( b(\cdot, \cdot) \) to these linear subspaces.

Let \( V \) be any vector space over \( k \) again, and suppose that \( b(\cdot, \cdot) \) is symmetric on \( V \). Let \( x \) be an element of \( V \) such that

\[
b(x, x) \neq 0,
\]

and let \( V_0 \) be the linear span of \( x \) in \( V \). Clearly \( V_0 \) is a one-dimensional linear subspace of \( V \), and the restriction of \( b(\cdot, \cdot) \) to \( V_0 \) is nondegenerate on \( V_0 \). This leads to a complementary subspace \( Z_0 \) of \( V_0 \) in \( V \) as before, so that \( b(\cdot, \cdot) \) corresponds to the bilinear form on the direct sum obtained from the restrictions of \( b(\cdot, \cdot) \) to \( V_0 \) and \( Z_0 \).

Suppose that \( V \) has finite dimension as a vector space over \( k \) again. Repeating the argument from the preceding paragraph, we get that \( V \) can be expressed as the direct sum of finitely many one-dimensional linear subspaces on which \( b(\cdot, \cdot) \neq 0 \), and possibly an additional linear subspace \( W \) such that (11.12.12) holds for every \( w \in W \). As before, \( b(\cdot, \cdot) \) corresponds to the bilinear form on the direct sum obtained from the restrictions of \( b(\cdot, \cdot) \) to these linear subspaces. If the characteristic of \( k \) is different from 2, then \( b(\cdot, \cdot) \equiv 0 \) on \( W \), because \( b(\cdot, \cdot) \) is both symmetric and antisymmetric on \( W \). In this case, the additional subspace \( W \) is not needed when \( b(\cdot, \cdot) \) is nondegenerate on \( V \).

### 11.13 Antisymmetric bilinear forms

Let \( k \) be a field with characteristic different from 2, and let \( V \) be a vector space over \( k \) of positive finite dimension. Also let \( \beta(\cdot, \cdot) \) be a nondegenerate antisymmetric bilinear form on \( V \). Under these conditions, there is a basis for \( V \) consisting of vectors \( x_1, \ldots, x_n, y_1, \ldots, y_n \) for some positive integer \( n \) with the following properties. First,

\[
\beta(x_j, x_l) = \beta(y_j, y_l) = 0
\]

for every \( j, l = 1, \ldots, n \), and

\[
\beta(x_j, y_l) = 0
\]
when \( j \neq l \). Second, for each \( j = 1, \ldots, n \),
\[
\beta(x_j, y_j) \neq 0.
\]

This well-known representation for \( \beta(\cdot, \cdot) \) follows from some of the remarks in
the previous section. More precisely, we may ask that
\[
\beta(x_j, y_j) = 1
\]
for every \( j = 1, \ldots, n \), by multiplying the \( x_j \)'s or \( y_j \)'s by suitable nonzero ele-
ments of \( k \), if necessary.

If \( v \in V \), then \( v \) can be expressed in a unique way as
\[
v = \sum_{j=1}^{n} v_{x_j} x_j + \sum_{j=1}^{n} v_{y_j} y_j,
\]
where \( v_{x_j}, v_{y_j} \in k \) for every \( j = 1, \ldots, n \). Put
\[
\langle v, w \rangle_V = \sum_{j=1}^{n} v_{x_j} w_{x_j} + \sum_{j=1}^{n} v_{y_j} w_{y_j}
\]
for every \( v, w \in V \), where \( w_{x_j}, w_{y_j} \in k \) correspond to \( w \) as in (11.13.5). This
defines a symmetric bilinear form on \( V \), with
\[
\langle v, v \rangle_V = \sum_{j=1}^{n} v_{x_j}^2 + \sum_{j=1}^{n} v_{y_j}^2
\]
for every \( v \in V \). If \( k \) is an ordered field, then (11.13.7) is strictly positive
when \( v \neq 0 \), and hence (11.13.6) defines an inner product on \( V \). Of course,
\( x_1, \ldots, x_n, y_1, \ldots, y_n \) is an orthonormal basis for \( V \) with respect to (11.13.6).

Let \( B \) be the unique linear mapping from \( V \) into itself that satis-
\[
B(x_j) = y_j, \quad B(y_j) = -x_j
\]
for every \( j = 1, \ldots, n \). It is easy to see that
\[
B^2 = -I_V,
\]
where \( I_V \) is the identity mapping on \( V \). One can check that
\[
\beta(v, w) = \langle B(v), w \rangle_V
\]
for every \( v, w \in V \), by reducing to the cases where \( v = x_j \) or \( y_j \) and \( w = x_l \) or
\( y_l \), \( 1 \leq j, l \leq n \). We also have that
\[
B^* = -B,
\]
where \( B^* \) is the adjoint of \( B \) with respect to (11.13.6). This corresponds to the
antisymmetry of \( \beta(\cdot, \cdot) \) on \( V \), because of (11.13.10).
If $T$ is a linear mapping from $V$ into itself, then the adjoint $T^{*,\beta}$ of $T$ with respect to $\beta(\cdot, \cdot)$ is the unique linear mapping from $V$ into itself such that
\begin{equation}
\beta(T(v), w) = \beta(v, T^{*,\beta}(w))
\end{equation}
for every $v, w \in V$. This is the same as saying that
\begin{equation}
\langle B(T(v)), w \rangle_V = \langle B(v), T^{*,\beta}(w) \rangle_V
\end{equation}
for every $v, w \in V$, by (11.13.10). This is equivalent to
\begin{equation}
\langle B(T(v)), w \rangle_V = -\langle v, B(T^{*,\beta}(w)) \rangle_V
\end{equation}
for every $v, w \in V$, because of (11.13.11). This means that
\begin{equation}
(B \circ T)^* = -B \circ T^{*,\beta},
\end{equation}
where the left side is the adjoint of $B \circ T$ with respect to the inner product (11.13.6). Thus
\begin{equation}
T^{*,\beta} = -B^{-1} \circ (B \circ T)^* = -B^{-1} \circ T^* \circ B = B^{-1} \circ T^* \circ B.
\end{equation}

### 11.14 Symmetric bilinear forms

Let $k$ be an ordered field, and let $V$ be a vector space over $k$ of positive finite dimension $n$. In this section, we consider a nondegenerate symmetric bilinear form $\beta(\cdot, \cdot)$ on $V$. It is well known that there is a basis $e_1, \ldots, e_n$ for $V$ such that
\begin{equation}
\beta(e_j, e_l) = 0
\end{equation}
when $j \neq l$, and for each $j = 1, \ldots, n$,
\begin{equation}
\beta(e_j, e_j) \neq 0,
\end{equation}
as in Section 11.12. Remember that the absolute value $|t|$ of $t \in k$ can be defined as an element of $k$, as in Section 8.13. If $k = \mathbb{R}$, then one can choose the $e_j$’s so that $\beta(e_j, e_j) = \pm 1$ for each $j = 1, \ldots, n$.

If $v \in V$, then let
\begin{equation}
v = \sum_{j=1}^{n} v_j e_j
\end{equation}
be the unique representation of $v$ as a linear combination of the $e_j$’s with coefficients $v_j \in k$. Put
\begin{equation}
\langle v, w \rangle_V = \sum_{j=1}^{n} |\beta(e_j, e_j)| v_j w_j
\end{equation}
for every $v, w \in V$, where $w_j \in k$ corresponds to $w$ as in (11.14.3). This defines a symmetric bilinear form on $V$, and
\begin{equation}
\langle v, v \rangle_V = \sum_{j=1}^{n} |\beta(e_j, e_j)| v_j^2
\end{equation}
for every $v \in V$. If $v \neq 0$, then (11.14.5) is strictly positive in $k$, so that (11.14.4) is an inner product on $V$. Note that the $e_j$’s are pairwise orthogonal with respect to (11.14.4).

Let $B$ be the unique linear mapping from $V$ into itself that satisfies

\begin{equation}
B(e_j) = \begin{cases} 
  e_j & \text{when } \beta(e_j, e_j) > 0 \\
  -e_j & \text{when } -\beta(e_j, e_j) > 0.
\end{cases}
\end{equation}

Clearly

\begin{equation}
B^2 = I_V,
\end{equation}
where $I_V$ is the identity mapping on $V$. It is easy to see that

\begin{equation}
B^* = B,
\end{equation}
where $B^*$ is the adjoint of $B$ with respect to the inner product (11.14.4). Of course,

\begin{equation}
\langle B(v), w \rangle_V = \sum_{j=1}^{n} \beta(e_j, e_j) v_j w_j
\end{equation}
for every $v, w \in V$, by construction. One can check that

\begin{equation}
\beta(v, w) = \langle B(v), w \rangle_V
\end{equation}
for every $v, w \in V$, by reducing to the case where $v = e_j$ and $w = e_l, 1 \leq j, l \leq n$.

If $T$ is a linear mapping from $V$ into itself, then the adjoint $T^{*, \beta}$ of $T$ with respect to $\beta$ is the unique linear mapping from $V$ into itself such that

\begin{equation}
\beta(T(v), w) = \beta(v, T^{*, \beta}(w))
\end{equation}
for every $v, w \in V$, as before. This means that

\begin{equation}
\langle B(T(v)), w \rangle_V = \langle B(v), T^{*, \beta}(w) \rangle_V
\end{equation}
for every $v, w \in V$, because of (11.14.10). This is the same as saying that

\begin{equation}
\langle B(T(v)), w \rangle_V = \langle v, B(T^{*, \beta}(w)) \rangle_V
\end{equation}
for every $v, w \in V$, by (11.14.8). This is equivalent to asking that

\begin{equation}
(B \circ T)^* = B \circ T^{*, \beta},
\end{equation}
where the left side is the adjoint of $B \circ T$ with respect to the inner product (11.14.4). This shows that

\begin{equation}
T^{*, \beta} = B^{-1} \circ (B \circ T)^* = B^{-1} \circ T^* \circ B^* = B^{-1} \circ T^* \circ B.
\end{equation}
Chapter 12

Some complex versions

12.1 Complexifying ordered fields

Let \( k \) be an ordered ring, as in Section 8.12, and suppose that \( k \) is commutative and has a multiplicative identity element \( 1 = 1_k \). Also let \( a \) be an element of \( k \) such that \( a > 0 \). Thus, for each \( x \in k \), we have that \( x^2 \neq -a \), because \( x^2 \geq 0 \). As usual, we can get a commutative ring \( k[\sqrt{-a}] \) by adjoining a square root \( \sqrt{-a} \) of \( -a \) to \( k \). More precisely, we can define \( k[\sqrt{-a}] \) to be the space \( k^2 \) of ordered pairs of elements of \( k \). An element of \( k[\sqrt{-a}] \) may be expressed in a unique way as

\[
(12.1.1) \quad z = x + y \sqrt{-a},
\]

with \( x, y \in k \), which corresponds to \((x, y) \in k^2\). Let us identify \( x \in k \) with \( x + 0 \sqrt{-a} \) in \( k[\sqrt{-a}] \), so that \( k \) corresponds to a subset of \( k[\sqrt{-a}] \). Let

\[
(12.1.2) \quad w = u + v \sqrt{-a}
\]

be another element of \( k[\sqrt{-a}] \), with \( u, v \in k \). Addition and multiplication on \( k \) can be extended to \( k[\sqrt{-a}] \) by putting

\[
(12.1.3) \quad z + w = (x + u) + (y + v) \sqrt{-a}
\]

and

\[
(12.1.4) \quad z w = (x u - y v a) + (x v + y u) \sqrt{-a},
\]

as usual. One can verify that \( k[\sqrt{-a}] \) is a commutative ring, and that \( k \) corresponds to a subring of \( k[\sqrt{-a}] \). The multiplicative identity element 1 in \( k \) corresponds to the multiplicative identity element in \( k[\sqrt{-a}] \). Of course, \((\sqrt{-a})^2 = -a \) in \( k[\sqrt{-a}] \), by construction.

If \( z \in k[\sqrt{-a}] \) is as in (12.1.1), then the conjugate \( \overline{z} \) of \( z \) in \( k[\sqrt{-a}] \) is defined as usual by

\[
(12.1.5) \quad \overline{z} = x - y \sqrt{-a}.
\]

One can verify that

\[
(12.1.6) \quad \overline{z} + w = \overline{z} + \overline{w}
\]
and
\[(12.1.7) \quad \overline{zw} = \overline{z} \overline{w}\]
for every \(z, w \in k[\sqrt{-a}]\). Clearly
\[(12.1.8) \quad \overline{z} = z\]
for every \(z \in k[\sqrt{-a}]\), and \(\overline{z} = z\) exactly when \(z\) corresponds to an element of \(k\). If \(z \in k[\sqrt{-a}]\) is as in (12.1.1) again, then
\[(12.1.9) \quad z + \overline{z} = 2 \cdot x\]
and
\[(12.1.10) \quad z \overline{z} = x^2 + ay^2\]
correspond to elements of \(k\). Note that
\[(12.1.11) \quad z \overline{z} > 0\]
in \(k\) when \(z \neq 0\).

Suppose now that \(k\) is an ordered field. If \(z \in k[\sqrt{-a}]\) and \(z \neq 0\), then \(z \overline{z}\) corresponds to a nonzero element of \(k\), which has a multiplicative inverse in \(k\). This implies that \(z\) has a multiplicative inverse in \(k[\sqrt{-a}]\), which is \(\overline{z}(1/(z \overline{z}))\). Thus \(k[\sqrt{-a}]\) is a field in this case as well. If \(k\) is a subfield of \(\mathbb{R}\), then we can take \(\sqrt{-a}\) to be \(i \sqrt{a} \in \mathbb{C}\), so that \(k[\sqrt{-a}]\) corresponds to a subfield of \(\mathbb{C}\).

Let \(k\) be an ordered field again, let \(a\) be a positive element of \(k\), and let \(k[\sqrt{-a}]\) be as before. Also let \(V\) and \(W\) be vector spaces over \(k[\sqrt{-a}]\), which may be considered as vector spaces over \(k\) too. A linear mapping \(T\) from \(V\) into \(W\), as vector spaces over \(k[\sqrt{-a}]\), may be called \(k[\sqrt{-a}]\)-linear, or linear over \(k[\sqrt{-a}]\), as usual. Similarly, if \(T\) is a linear mapping from \(V\) into \(W\) as vector spaces over \(k\), the we may say that \(T\) is \(k\)-linear, or linear over \(k\). Thus a mapping \(T\) from \(V\) into \(W\) is \(k[\sqrt{-a}]\)-linear if and only if \(T\) is \(k\)-linear and
\[(12.1.12) \quad T(\sqrt{-a} v) = \sqrt{-a} T(v)\]
for every \(v \in V\). A \(k\)-linear mapping \(T\) from \(V\) into \(W\) is said to be *conjugate-linear* if
\[(12.1.13) \quad T(\sqrt{-a} v) = -\sqrt{-a} T(v)\]
for every \(v \in V\). This implies that
\[(12.1.14) \quad T(t v) = t T(v)\]
for every \(t \in k[\sqrt{-a}]\) and \(v \in V\).

### 12.2 Sesquilinearity over \(k[\sqrt{-a}]\)

Let \(k\) be an ordered field, let \(a\) be a positive element of \(k\), and let \(k[\sqrt{-a}]\) be as in the previous section. Also let \(V\) be a vector space over \(k\), and let \(b(v, w)\)
be a function defined for \( v, w \in V \) with values in \( k[\sqrt{-a}] \). We say that \( b(\cdot, \cdot) \) is sesquilinear if \( b(v, w) = k[\sqrt{-a}] \)-linear in \( v \) for every \( v \in V \), and conjugate-linear in \( w \) for every \( v \in V \). In particular, this means that \( b(\cdot, \cdot) \) is bilinear over \( k \), where \( V \) and \( k[\sqrt{-a}] \) are considered as vector spaces over \( k \). As before, \( b(\cdot, \cdot) \) is said to be Hermitian-symmetric on \( V \) if

\[
(12.2.1) \quad b(w, v) = \overline{b(v, w)}
\]

for every \( v, w \in V \).

Let \( b(\cdot, \cdot) \) be a sesquilinear form on \( V \), and let \( T \) be a \( k[\sqrt{-a}] \)-linear mapping from \( V \) into itself. As before, we say that \( T \) is self-adjoint with respect to \( b(\cdot, \cdot) \) if

\[
(12.2.2) \quad b(T(v), w) = b(v, T(w))
\]

for every \( v, w \in V \), and we say that \( T \) is anti-self-adjoint with respect to \( b(\cdot, \cdot) \) if

\[
(12.2.3) \quad b(T(v), w) = -b(v, T(w))
\]

for every \( v, w \in V \). It is easy to see that \( T \) is anti-self-adjoint with respect to \( b(\cdot, \cdot) \) if and only if \( \sqrt{-a} T \) is self-adjoint with respect to \( b(\cdot, \cdot) \). Remember that the space \( \mathcal{L}(V) \) of \( k[\sqrt{-a}] \)-linear mappings from \( V \) into itself is a vector space over \( k[\sqrt{-a}] \), and thus may be considered as a vector space over \( k \) too. The spaces of self-adjoint and anti-self-adjoint linear mappings from \( V \) into itself with respect to \( b(\cdot, \cdot) \) are \( k \)-linear subspaces of \( \mathcal{L}(V) \), which is to say that they are linear subspaces of \( \mathcal{L}(V) \), as a vector space over \( k \). If \( T_1, T_2 \in \mathcal{L}(V) \) are anti-self-adjoint with respect to \( b(\cdot, \cdot) \), then their commutator \( [T_1, T_2] = T_1 \circ T_2 - T_2 \circ T_1 \) with respect to composition is anti-self-adjoint with respect to \( b(\cdot, \cdot) \) as well. This means that the space of anti-self-adjoint linear mappings from \( V \) into itself with respect to \( b(\cdot, \cdot) \) is a Lie subalgebra of \( gl(V) \) as a Lie algebra over \( k \) with respect to the commutator bracket.

Suppose from now on in this section that \( V \) has finite dimension as a vector space over \( k[\sqrt{-a}] \). If \( w \in V \), then

\[
(12.2.4) \quad b_w(v) = b(v, w)
\]

defines a linear functional on \( V \), and \( w \mapsto b_w \) is a conjugate-linear mapping from \( V \) into its dual space \( V' \). The image

\[
(12.2.5) \quad \{b_w : w \in V\}
\]

of this mapping is a linear subspace of \( V' \), as a vector space over \( k[\sqrt{-a}] \). As before, one can verify that \( (12.2.5) \) is equal to \( V' \) exactly when \( w \mapsto b_w \) is injective, because \( V \) and \( V' \) have the same dimension as vector spaces over \( k[\sqrt{-a}] \).

We say that \( b(\cdot, \cdot) \) is nondegenerate as a sesquilinear form on \( V \) if for every \( v \in V \) with \( v \neq 0 \) there is a \( w \in V \) such that \( b(v, w) \neq 0 \). Equivalently, this means that the intersections of the kernels of the \( b_w \)'s, \( w \in V \), is the trivial subspace of \( V \). This holds exactly when \( (12.2.5) \) is equal to \( V' \), as in the complex case.
12.3. INNER PRODUCTS OVER $K[\sqrt{-A}]$

Let $b(\cdot, \cdot)$ be a nondegenerate sesquilinear form on $V$, let $T$ be a $k[\sqrt{-a}]$-linear mapping from $V$ into itself, and let $w \in V$ be given. Thus $b(T(v), w)$ defines a linear functional on $V$, as a function of $v$, so that there is a unique element $T^*(w)$ of $V$ such that

\[(12.2.6) \quad b(T(v), w) = b(v, T^*(w))\]

for every $v \in V$. One can check that $T^*$ is a $k[\sqrt{-a}]$-linear mapping from $V$ into itself, which is called the adjoint of $T$ with respect to $b(\cdot, \cdot)$. However, $T \mapsto T^*$ is conjugate-linear as a mapping from $L(V)$ into itself, as a vector space over $k[\sqrt{-a}]$. Note that $T$ is self-adjoint with respect to $b(\cdot, \cdot)$ when $T = T^*$, and that $T$ is anti-self-adjoint with respect to $b(\cdot, \cdot)$ when $T^* = -T$. If $T_1, T_2 \in L(V)$, then one can check that

\[(12.2.7) \quad (T_2 \circ T_1)^* = T_1^* \circ T_2^*, \]

as usual. If $b(\cdot, \cdot)$ is Hermitian-symmetric on $V$, then one can verify that

\[(12.2.8) \quad (T^*)^* = T\]

for every $T \in L(V)$, as in Section 2.15.

12.3 Inner products over $k[\sqrt{-a}]$

Let $k$ be an ordered field again, let $a$ be a positive element of $k$, and let $k[\sqrt{-a}]$ be as in Section 12.1. Also let $V$ be a vector space over $k[\sqrt{-a}]$, and let $\langle \cdot, \cdot \rangle_V$ be a Hermitian form on $V$, which is to say a Hermitian-symmetric sesquilinear form on $V$. Note that

\[(12.3.1) \quad \langle v, v \rangle_V \in k\]

for every $v \in V$, because $\overline{\langle v, v \rangle_V} = \langle v, v \rangle_V$, by Hermitian symmetry. Suppose that

\[(12.3.2) \quad \langle v, v \rangle_V > 0\]

for every $v \in V$ with $v \neq 0$, in which case $\langle \cdot, \cdot \rangle_V$ is said to be an inner product on $V$. If $v, w \in V$ satisfy

\[(12.3.3) \quad \langle v, w \rangle_V = 0,\]

then $v$ and $w$ are said to be orthogonal with respect to $\langle \cdot, \cdot \rangle_V$, as usual.

Let $e_1, \ldots, e_n$ be finitely many pairwise-orthogonal nonzero vectors in $V$, and let $v \in V$ be given. Thus

\[(12.3.4) \quad w = \sum_{j=1}^{n} \frac{\langle v, e_j \rangle_V}{\langle e_j, e_j \rangle_V} e_j\]

is an element of the linear span of $e_1, \ldots, e_n$ in $V$, and

\[(12.3.5) \quad \langle w, e_l \rangle_V = \langle v, e_l \rangle_V\]
for every $l = 1, \ldots, n$. This means that $\langle w - v, e_l \rangle_V = 0$ for each $l = 1, \ldots, n$, so that

\[
\langle w - v, u \rangle_V = 0
\]

for every element $u$ of the linear span of $e_1, \ldots, e_n$ in $V$. If $v$ is an element of the linear span of $e_1, \ldots, e_n$ in $V$, then it follows that $v = w$.

Let us suppose from now on in this section that $V$ has positive finite dimension $n$, as a vector space over $k[\sqrt{-a}]$. Using the Gram–Schmidt process, one can get nonzero pairwise-orthogonal vectors $e_1, \ldots, e_n$ in $V$ that form a basis for $V$. If $v \in V$, then

\[
v = \sum_{j=1}^{n} \frac{\langle v, e_j \rangle_V}{\langle e_j, e_j \rangle_V} e_j,
\]

as in the preceding paragraph. Let $T$ be a linear mapping from $V$ into itself, so that there is a unique $T^* \in \mathcal{L}(V)$ such that

\[
\langle T(v), w \rangle_V = \langle v, T^*(w) \rangle_V
\]

for every $v, w \in V$, as before. Of course,

\[
T(e_l) = \sum_{j=1}^{n} \frac{\langle T(e_l), e_j \rangle_V}{\langle e_j, e_j \rangle_V} e_j
\]

for every $l = 1, \ldots, n$, as in (12.3.7). This implies that

\[
\text{tr}_V T = \sum_{j=1}^{n} \frac{\langle T(e_j), e_j \rangle_V}{\langle e_j, e_j \rangle_V}.
\]

Thus

\[
\text{tr}_V T^* = \sum_{j=1}^{n} \frac{\langle T^*(e_j), e_j \rangle_V}{\langle e_j, e_j \rangle_V} = \sum_{j=1}^{n} \frac{\langle e_j, T(e_j) \rangle_V}{\langle e_j, e_j \rangle_V} = \sum_{j=1}^{n} \frac{\langle T(e_j), e_j \rangle_V}{\langle e_j, e_j \rangle_V} = \text{tr}_V T.
\]

Put

\[
\langle T_1, T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (T_1 \circ T_2^*) = \text{tr}_V (T_2^* \circ T_1)
\]

for every $T_1, T_2 \in \mathcal{L}(V)$. This defines a sesquilinear form on $\mathcal{L}(V)$, because $T \mapsto T^*$ is conjugate-linear on $\mathcal{L}(V)$. Observe that

\[
\langle T_1, T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (T_1 \circ T_2^*) = \text{tr}_V (T_2^* \circ T_1)
\]

(12.3.13)
12.4. SELF-ADJOINTNESS IN GL(V)

for every $T_1, T_2 \in \mathcal{L}(V)$, using (12.3.11) in the second step. Thus (12.3.12) is Hermitian-symmetric on $\mathcal{L}(V)$. If $T \in \mathcal{L}(V)$, then

\begin{equation}
\langle T, T \rangle_{\mathcal{L}(V)} = \text{tr}_V (T^* \circ T) = \sum_{j=1}^{n} \frac{\langle T^*(e_j), e_j \rangle_V}{\langle e_j, e_j \rangle_V} = \sum_{j=1}^{n} \frac{\langle T(e_j), e_j \rangle_V}{\langle e_j, e_j \rangle_V}.
\end{equation}

This implies that $\langle T, T \rangle_{\mathcal{L}(V)} > 0$ when $T \neq 0$, because each of the terms on the right side of (12.3.14) is greater than or equal to 0 in $k$, and at least one term is strictly positive. It follows that (12.3.12) defines an inner product on $\mathcal{L}(V)$, as a vector space over $k[\sqrt{-a}]$.

Let $A \in \mathcal{L}(V)$ be given, and put

\begin{equation}
L_A(T) = A \circ T, \quad R_A(T) = T \circ A
\end{equation}

for every $T \in \mathcal{L}(V)$, as before. As in Sections 11.8 and 11.9, we have that

\begin{equation}
\langle L_A(T_1), T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (A \circ T_1 \circ T_2) = \text{tr}_V (T_1 \circ T_2^* \circ A) = \langle T_1, L_{A^*}(T_2) \rangle_{\mathcal{L}(V)}
\end{equation}

and

\begin{equation}
\langle R_A(T_1), T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (T_1 \circ A \circ T_2^*) = \text{tr}_V (T_1 \circ (T_2 \circ A)^*) = \langle T_1, R_{A^*}(T_2) \rangle_{\mathcal{L}(V)}
\end{equation}

for every $T_1, T_2 \in \mathcal{L}(V)$. If we put

\begin{equation}
C_A(T) = [A, T] = L_A(T) - R_A(T)
\end{equation}

for each $T \in \mathcal{L}(V)$, then we get that

\begin{equation}
\langle C_A(T_1), T_2 \rangle_{\mathcal{L}(V)} = \langle T_1, C_{A^*}(T_2) \rangle_{\mathcal{L}(V)}
\end{equation}

for every $T_1, T_2 \in \mathcal{L}(V)$, as before.

12.4  Self-adjointness in $gl(V)$

Let us continue with the same notation and hypotheses as in the previous section. We would like to consider the analogues of the remarks in Section 11.10 in this situation. Remember that $gl(V)$ is the same as $\mathcal{L}(V)$, considered as a Lie algebra over $k[\sqrt{-a}]$ with respect to the corresponding commutator bracket. If $R, T \in gl(V)$, then

\begin{equation}
C_R(T) = [R, T] = R \circ T - T \circ R
\end{equation}

is the same as $\text{ad}_{gl(V), R}(T)$. As before,

\begin{equation}
\langle C_R(T_1), T_2 \rangle_{\mathcal{L}(V)} = \langle T_1, C_{R^*}(T_2) \rangle_{\mathcal{L}(V)}
\end{equation}
for every $R, T_1, T_2 \in gl(V)$.

Let $\mathcal{A}$ be a Lie subalgebra of $gl(V)$, as a Lie algebra over $k[\sqrt{-a}]$. Thus, for

\textit{the adjoint of $ad$}

each $R \in \mathcal{A}$, $C_R$ maps $\mathcal{A}$ into itself, and the restriction of $C_R$ to $\mathcal{A}$ is the same

as $\text{ad}_{A,R}$. Let $R_1, R_2 \in \mathcal{A}$ be given, so that

\begin{equation}
(12.4.3) \quad \text{tr}_{\mathcal{A}}(\text{ad}_{A,R_1} \circ \text{ad}_{A,R_2}) = \text{tr}_{\mathcal{A}}(C_{R_1} \circ C_{R_2})
\end{equation}

is the same as the Killing form on $\mathcal{A}$ evaluated at $R_1, R_2$, as before. This uses the trace on $\mathcal{A}$, as a finite-dimensional vector space over $k[\sqrt{-a}]$. More precisely, one should use the restrictions of $C_{R_1}$ and $C_{R_2}$ to $\mathcal{A}$ on the right side of (12.4.3).

If $\mathcal{A}$ is any subset of $L(V)$, then let $\mathcal{A}^*$ be the subset of $L(V)$ defined by

\begin{equation}
(12.4.4) \quad \mathcal{A}^* = \{ T^* : T \in \mathcal{A} \},
\end{equation}

as before. Let us say that $\mathcal{A}$ is self-adjoint in $L(V)$ when $\mathcal{A}^* = \mathcal{A}$. Let $\mathcal{A}$ be a

Lie subalgebra of $gl(V)$ again, as a Lie algebra over $k[\sqrt{-a}]$, and suppose that $\mathcal{A}$ is self-adjoint in $L(V)$. Let $R \in \mathcal{A}$ be given, so that $R^* \in \mathcal{A}$ too, and $C_R$ and $C_{R^*}$ both map $\mathcal{A}$ into itself. The restriction of $C_{R^*}$ to $\mathcal{A}$ is the adjoint of the restriction of $C_R$ to $\mathcal{A}$, as an inner product space over $k[\sqrt{-a}]$ with respect to the restriction of the inner product (12.3.12) on $L(V)$ to $\mathcal{A}$.

Let $L(\mathcal{A})$ be the space of all linear mappings from $\mathcal{A}$ into itself, which is

in particular a vector space over $k[\sqrt{-a}]$. The restriction of the inner product (12.3.12) on $L(V)$ to $\mathcal{A}$ defines an inner product on $\mathcal{A}$ as a vector space over $k[\sqrt{-a}]$, as before. If $Z \in L(\mathcal{A})$, then let $Z_{A^*} \in L(\mathcal{A})$ be the adjoint of $Z$ with respect to this inner product. If $Z_1, Z_2 \in L(\mathcal{A})$, then put

\begin{equation}
(12.4.5) \quad \langle Z_1, Z_2 \rangle_{L(\mathcal{A})} = \text{tr}_{\mathcal{A}}(Z_1 \circ Z_{2 A^*}).
\end{equation}

This defines an inner product on $L(\mathcal{A})$, as a vector space over $k[\sqrt{-a}]$, as in the

previous section.

Let $R_1, R_2 \in \mathcal{A}$ be given, so that $R^*_2 \in \mathcal{A}$ as well, and $\text{ad}_{A,R_2}$ is the same as

the adjoint of $\text{ad}_{A,R_2}$ with respect to the inner product on $\mathcal{A}$ mentioned in the

preceding paragraph. This implies that

\begin{equation}
(12.4.6) \quad \text{tr}_{\mathcal{A}}(\text{ad}_{A,R_1} \circ \text{ad}_{A,R_2}) = \langle \text{ad}_{A,R_1}, \text{ad}_{A,R_2} \rangle_{L(\mathcal{A})},
\end{equation}

where the right side is as in (12.4.5). If $R \in \mathcal{A}$, then we can take $R_1 = R$ and

$R_2 = R^*$, to get that

\begin{equation}
(12.4.7) \quad \text{tr}_{\mathcal{A}}(\text{ad}_{A,R} \circ \text{ad}_{A,R^*}) = \langle \text{ad}_{A,R}, \text{ad}_{A,R^*} \rangle_{L(\mathcal{A})}.
\end{equation}

The right side is automatically greater than or equal to 0 in $k$, because (12.4.5) is an inner product on $L(\mathcal{A})$. Similarly, the right side is strictly positive when $\text{ad}_{A,R} \neq 0$ on $\mathcal{A}$. 
12.5 Projections and sesquilinear forms

Let $k$ be an ordered field, let $a$ be a positive element of $k$ again, and let $k[\sqrt{-a}]$ be as in Section 12.1. Also let $V$ be a vector space over the complex numbers, and let $b(v, w)$ be a sesquilinear form on $V$, as in Section 12.2. Suppose that $V_0$ is a finite-dimensional linear subspace of $V$, and that the restriction of $b(v, w)$ to $v, w \in V_0$ is nondegenerate on $V_0$. Let $z \in V$ be given, and put

$$\lambda_z(v) = b(v, z)$$

for every $v \in V_0$, so that $\lambda_z$ defines a linear functional on $V_0$. As in Section 11.12, there is a unique element $P_0(z)$ of $V_0$ such that

$$\lambda_z(v) = b(v, P_0(z))$$

for every $v \in V_0$, because $b(\cdot, \cdot)$ is nondegenerate on $V_0$. Equivalently,

$$b(v, z - P_0(z)) = 0$$

for every $v \in V$. Of course,

$$P_0(z) = z$$

when $z \in V_0$. One can check that $P_0$ defines a linear mapping from $V$ onto $V_0$, as before. Note that $P_0 \circ P_0 = P_0$, so that $P_0$ is a projection on $V$.

Let $Z_0$ be the kernel of $P_0$, and observe that

$$V_0 \cap Z_0 = \{0\},$$

by (12.5.4). We also have that $z - P_0(z) \in Z_0$ for every $z \in V$, so that

$$V = V_0 + Z_0.$$

Thus $V$ corresponds to the direct sum of $V_0$ and $Z_0$, as a vector space over $k[\sqrt{-a}]$. Note that $Z_0$ is the same as the set of $z \in V$ such that

$$b(v, z) = 0$$

for every $v \in V_0$.

Suppose from now on in this section that $b(\cdot, \cdot)$ is Hermitian-symmetric on $V$. Using this and (12.5.7), we get that

$$b(z, v) = 0$$

for every $v \in V_0$ and $z \in Z_0$. Let $u, w \in V$ be given, so that $P_0(u), P_0(w) \in V_0$, $u - P_0(u), w - P_0(w) \in Z_0$, and

$$b(P_0(u), w - P_0(w)) = b(u - P_0(u), P_0(w)) = 0.$$

Thus

$$b(u, w) = b(P_0(u), P_0(w)) + b(u - P_0(u), w - P_0(w)).$$
This means that $b(\cdot, \cdot)$ corresponds to the sesquilinear form on the direct sum of $V_0$ and $Z_0$ obtained from the restrictions of $b(\cdot, \cdot)$ to $V_0$ and $Z_0$.

Suppose that $x \in V$ satisfies

$$b(x, x) \neq 0,$$

and let $V_0$ be the linear span of $x$ in $V$. Thus $V_0$ is a one-dimensional linear subspace of $V$, and the restriction of $b(\cdot, \cdot)$ to $V_0$ is nondegenerate on $V_0$. It follows that $V$ corresponds to the direct sum of $V_0$ and another linear subspace $Z_0$, in such a way that $b(\cdot, \cdot)$ corresponds to the sesquilinear form on the direct sum obtained from the restrictions of $b(\cdot, \cdot)$ to $V_0$ and $Z_0$.

Suppose now that $V$ has finite dimension as a vector space over $k[\sqrt{-a}]$. Repeating the previous argument, we can express $V$ as the direct sum of finitely many one-dimensional subspaces on which $b(\cdot, \cdot) \neq 0$, and perhaps an additional linear subspace $W$ such that

$$b(w, w) = 0$$

for every $w \in W$. We also have that $b(\cdot, \cdot)$ corresponds to the sesquilinear form on the direct sum obtained from the restrictions of $b(\cdot, \cdot)$ to these linear subspaces. One can check that $b(\cdot, \cdot) \equiv 0$ on $W$, using polarization arguments. If $b(\cdot, \cdot)$ is nondegenerate on $V$, then this additional subspace $W$ is not needed.

12.6 Nondegenerate Hermitian forms

Let $k$ be an ordered field, let $a$ be a positive element of $k$, and let $k[\sqrt{-a}]$ be as in Section 12.1 again. Also let $V$ be a vector space over $k[\sqrt{-a}]$ of positive finite dimension $n$, and let $\beta(\cdot, \cdot)$ be a nondegenerate Hermitian form on $V$. Under these conditions, there is a basis $e_1, \ldots, e_n$ for $V$ such that

$$\beta(e_j, e_l) = 0$$

when $j \neq l$, and for each $j = 1, \ldots, n$,

$$\beta(e_j, e_j) \neq 0,$$

as in the previous section. Note that

$$\beta(e_j, e_j) \in k$$

for each $j = 1, \ldots, n$, by Hermitian symmetry. If $k = \mathbb{R}$, then one choose the $e_j$’s so that $\beta(e_j, e_j) = \pm 1$ for each $j = 1, \ldots, n$.

Each $v \in V$ can be expressed in a unique way as

$$v = \sum_{j=1}^{n} v_j e_j,$$
where \( v_j \in k[\sqrt{-a}] \) for every \( j = 1, \ldots, n \). Remember that the absolute value \(|t|\) of \( t \in k \) is defined as an element of \( k \) as in Section 8.13. If \( v, w \in V \), then put

\[
\langle v, w \rangle_V = \sum_{j=1}^{n} |\beta(e_j, e_j)| v_j \overline{w}_j,
\]

(12.6.5)

where \( w_j \in k[\sqrt{-a}] \) corresponds to \( w \) as in (12.6.4). This defines a Hermitian form on \( V \), as a vector space over \( k[\sqrt{-a}] \). In particular,

\[
\langle v, v \rangle_V = \sum_{j=1}^{n} |\beta(e_j, e_j)| v_j \overline{v}_j
\]

(12.6.6)

is an element of \( k \) for each \( v \in V \). If \( v \neq 0 \), then it is easy to see that (12.6.6) is positive in \( k \). Thus (12.6.5) defines an inner product on \( V \), as a vector space over \( k[\sqrt{-a}] \). Of course, the \( e_j \)'s are orthogonal with respect to (12.6.5), by construction.

Let \( B \) be the unique linear mapping from \( V \) into itself such that

\[
B(e_j) = \begin{cases} 
  e_j & \text{when } \beta(e_j, e_j) > 0 \\
  -e_j & \text{when } -\beta(e_j, e_j) > 0.
\end{cases}
\]

(12.6.7)

Note that \( B^2 \) is the identity mapping \( I = I_V \) on \( V \). One can check that \( B \) is self-adjoint with respect to (12.6.5). It is easy to see that

\[
\beta(v, w) = \langle B(v), w \rangle_V
\]

(12.6.8)

when \( v = e_j \) and \( w = e_l, 1 \leq j, l \leq n \). This implies that (12.6.8) holds for every \( v, w \in V \).

Let \( T \) be a linear mapping from \( V \) into itself, and remember that the adjoint \( T^{\ast,\beta} \in \mathcal{L}(V) \) of \( T \) with respect to \( \beta(\cdot, \cdot) \) is characterized by the condition that

\[
\beta(T(v), w) = \beta(v, T^{\ast,\beta}(w))
\]

(12.6.9)

for every \( v, w \in V \). This is the same as saying that

\[
\langle B(T(v)), w \rangle_V = \langle B(v), T^{\ast,\beta}(w) \rangle_V
\]

(12.6.10)

for every \( v, w \in V \), by (12.6.8). Equivalently, this means that

\[
\langle B(T(v)), w \rangle_V = \langle v, B(T^{\ast,\beta}(w)) \rangle_V
\]

(12.6.11)

for every \( v, w \in V \), because \( B \) is self-adjoint with respect to (12.6.5). This is the same as asking that

\[
(B \circ T)^{\ast} = B \circ T^{\ast,\beta},
\]

(12.6.12)

where the left side is the adjoint of \( B \circ T \) with respect to the inner product (12.6.5). Thus

\[
T^{\ast,\beta} = B^{-1} \circ (B \circ T)^{\ast} = B^{-1} \circ T^{\ast} \circ B = B^{-1} \circ T^{\ast} \circ B.
\]

(12.6.13)
12.7 Real parts

Let $k$ be an ordered field, let $a$ be a positive element of $k$, and let $k[\sqrt{-a}]$ be as in Section 12.1, as before. If $z = x + y\sqrt{-a} \in k[\sqrt{-a}]$, with $x, y \in k$, then let us call

\[(12.7.1) \quad \Re_{k[\sqrt{-a}]} z = x = (1/2) (z + \bar{z})\]

the real part of $z$. Note that

\[(12.7.2) \quad \Re_{k[\sqrt{-a}]} (\sqrt{-a} z) = -a y = (\sqrt{-a}/2) (z - \bar{z}).\]

Let $V = V_{k[\sqrt{-a}]}$ be a finite-dimensional vector space over $k[\sqrt{-a}]$, and let $\langle \cdot, \cdot \rangle_{V_{k[\sqrt{-a}]}}$ be an inner product on $V$. We shall use $V_k$ to denote $V$ considered as a vector space over $k$. It is easy to see that

\[(12.7.3) \quad (v, w)_{V_k} = \Re_{k[\sqrt{-a}]} ((v, w)_{V_{k[\sqrt{-a}]}})\]

defines an inner product on $V_k$, as a vector space over $k$.

Let $\mathcal{L}(V) = \mathcal{L}_{k[\sqrt{-a}]}(V_{k[\sqrt{-a}]})$ be the algebra of $k[\sqrt{-a}]$-linear mappings from $V$ into itself, as usual, and let $\mathcal{L}_k(V_k)$ be the algebra of $k$-linear mappings from $V$ into itself. If $T \in \mathcal{L}_{k[\sqrt{-a}]}(V_{k[\sqrt{-a}]})$, then its adjoint with respect to $\langle \cdot, \cdot \rangle_{V_{k[\sqrt{-a}]}}$ is the unique $k[\sqrt{-a}]$-linear mapping $T^{*}_{V_{k[\sqrt{-a}]}}$ from $V$ into itself such that

\[(12.7.4) \quad (T(v), w)_{V_{k[\sqrt{-a}]}} = (v, T^{*}_{V_{k[\sqrt{-a}]}}(w))_{V_{k[\sqrt{-a}]}}\]

for every $v, w \in V$. Similarly, if $T \in \mathcal{L}_k(V_k)$, then its adjoint with respect to $(12.7.3)$ is the unique $k$-linear mapping $T^{*}_{V_k}$ from $V$ into itself such that

\[(12.7.5) \quad (T(v), w)_{V_k} = (v, T^{*}_{V_k}(w))_{V_k}\]

for every $v, w \in V$. If $T \in \mathcal{L}_{k[\sqrt{-a}]}(V_{k[\sqrt{-a}]})$, then $T$ may be considered as an element of $\mathcal{L}_k(V_k)$, and

\[(12.7.6) \quad T^{*}_{V_k} = T^{*}_{V_{k[\sqrt{-a}]}}.\]

Put

\[(12.7.7) \quad J_a(v) = \sqrt{-a} v\]

for every $v \in V$, which defines a $k[\sqrt{-a}]$-linear mapping from $V$ into itself. By construction,

\[(12.7.8) \quad J_a^2 = -a I,\]

where $I$ is the identity mapping on $V$, and

\[(12.7.9) \quad (J_a(v), w)_{V_k} = \Re_{k[\sqrt{-a}]} (\sqrt{-a} (v, w)_{V_{k[\sqrt{-a}]}})\]

for every $v, w \in V$. One can check that

\[(12.7.10) \quad J_a^{*}v_k = J_a^{*}v_{k[\sqrt{-a}]} = -J_a.\]
12.7. REAL PARTS

This means that (12.7.9) is an antisymmetric bilinear form on \( V_k \). Remember that a \( k \)-linear mapping from \( V \) into itself is linear over \( k[\sqrt{-a}] \) exactly when it commutes with \( J_a \).

Let \( \beta_{V_{k[\sqrt{-a}]}}(\cdot, \cdot) \) be a sesquilinear form on \( V_{k[\sqrt{-a}]} \), and put

\[
\beta_{V_k}(v, w) = \text{Re}_{k[\sqrt{-a}]}(\beta_{V_{k[\sqrt{-a}]}}(v, w))
\]

for every \( v, w \in V \). This defines a bilinear form on \( V_k \), which is to say a \( k \)-valued function on \( V \times V \) that is bilinear over \( k \). Observe that

\[
\beta_{V_k}(J_a(v), w) = -\beta_{V_k}(v, J_a(w)) = \text{Re}_{k[\sqrt{-a}]}(\sqrt{-a} \beta_{V_{k[\sqrt{-a}]}}(v, w))
\]

for every \( v, w \in V \). If \( \beta_{V_{k[\sqrt{-a}]}}(\cdot, \cdot) \) is Hermitian-symmetric on \( V_C \), then (12.7.11) is symmetric on \( V_{k_R} \) and (12.7.12) is antisymmetric on \( V_{k_R} \). Of course, (12.7.11) and (12.7.12) correspond to (12.7.3) and (12.7.9) when \( \beta_{V_C}(\cdot, \cdot) = (\cdot, \cdot)_{V_C} \).

Suppose now that \( \beta_{V_{k[\sqrt{-a}]}}(\cdot, \cdot) \) is nondegenerate as a sesquilinear form on \( V_{k[\sqrt{-a}]} \), which implies that (12.7.11) and (12.7.12) are nondegenerate as bilinear forms on \( V_k \). If \( T \in \mathcal{L}_{k[\sqrt{-a}]}(V_{k[\sqrt{-a}]}) \), then there is a unique \( T^{*,\beta_{V_{k[\sqrt{-a}]}}}_{k[\sqrt{-a}]} \) in \( \mathcal{L}_{k[\sqrt{-a}]}(V_{k[\sqrt{-a}]}) \) such that

\[
\beta_{V_{k[\sqrt{-a}]}}(T(v), w) = \beta_{V_{k[\sqrt{-a}]}}(v, T_{k[\sqrt{-a}]}^{*,\beta_{V_{k[\sqrt{-a}]}}}(w))
\]

for every \( v, w \in V \), as usual. Similarly, if \( T \in \mathcal{L}_k(V_k) \), then there is a unique \( T_{k[\sqrt{-a}]}^{*,\beta_{V_{k[\sqrt{-a}]}}}(V_k) \) such that

\[
\beta_{V_k}(T(v), w) = \beta_{V_k}(v, T_{k[\sqrt{-a}]}^{*,\beta_{V_{k[\sqrt{-a}]}}}(w))
\]

for every \( v, w \in V \). If \( T \in \mathcal{L}_{k[\sqrt{-a}]}(V_{k[\sqrt{-a}]}) \), then \( T \in \mathcal{L}_k(V_k) \), and

\[
T_{k[\sqrt{-a}]}^{*,\beta_{V_{k[\sqrt{-a}]}}}(V_k) = T^{*,\beta_{V_{k[\sqrt{-a}]}}} \quad \text{as before. In particular,} \quad J_{k[\sqrt{-a}]}^{*,\beta_{V_{k[\sqrt{-a}]}}} = -J_a.
\]

Let \( \alpha_{V_k}(\cdot, \cdot) \) be a bilinear form on \( V_k \) such that

\[
\alpha_{V_k}(J_a(v), w) = -\alpha_{V_k}(v, J_a(w))
\]

for every \( v, w \in V \). Put

\[
\alpha_{V_{k[\sqrt{-a}]}}(v, w) = \alpha_{V_k}(v, w) + (1/\sqrt{-a}) \alpha_{V_k}(J_a(v), w)
\]

for every \( v, w \in V \). One can check that this defines a sesquilinear form on \( V_{k[\sqrt{-a}]} \). If \( \alpha_{V_k}(\cdot, \cdot) \) is symmetric on \( V_k \), then (12.7.16) is antisymmetric on \( V_k \), and (12.7.17) is Hermitian-symmetric on \( V_{k[\sqrt{-a}]} \). If \( \alpha_{V_k}(\cdot, \cdot) \) is nondegenerate on \( V_k \), then (12.7.16) is nondegenerate on \( V_k \), and (12.7.17) is nondegenerate on \( V_{k[\sqrt{-a}]} \). If \( \alpha_{V_k}(\cdot, \cdot) \) is an inner product on \( V_k \), then (12.7.17) is an inner product on \( V_{k[\sqrt{-a}]} \). If \( \alpha_{V_k}(\cdot, \cdot) \) is equal to (12.7.11), then (12.7.17) is equal to \( \beta_{V_{k[\sqrt{-a}]}}(\cdot, \cdot) \), by (12.7.12).
12.8 Bilinear forms over $k[\sqrt{-a}]$

Let $k$ be an ordered field, let $a$ be a positive element of $k$, and let $k[\sqrt{-a}]$ be as in Section 12.1, as usual. Also let $V = V_{k[\sqrt{-a}]}$ be a finite-dimensional vector space over $k[\sqrt{-a}]$ again, and let $\gamma_{V_{k[\sqrt{-a}]}}(\cdot, \cdot)$ be a bilinear form on $V_{k[\sqrt{-a}]}$, which is to say a $k[\sqrt{-a}]$-valued function on $V \times V$ that is bilinear over $k[\sqrt{-a}]$. It follows that

\begin{equation}
\gamma_{V_{k}}(v, w) = \operatorname{Re}_{k[\sqrt{-a}]}(\gamma_{V_{k[\sqrt{-a}]}}(v, w))
\end{equation}

defines a bilinear form on $V_{k}$, which is to say a $k$-valued function on $V \times V$ that is bilinear over $k$. Put $J_{a}(v) = \sqrt{-a}v$ for every $v \in V$, as in the previous section. Observe that

\begin{equation}
\gamma_{V_{k}}(v, w) = \gamma_{V_{k}}(v, J_{a}(w)) = \operatorname{Re}_{k[\sqrt{-a}]}(\sqrt{-a} \gamma_{V_{k[\sqrt{-a}]}}(v, w))
\end{equation}

for every $v, w \in V$. If $\gamma_{V_{k[\sqrt{-a}]}}(\cdot, \cdot)$ is symmetric or antisymmetric on $V_{k[\sqrt{-a}]}$, then (12.8.1) and (12.8.2) have the same property on $V_{k}$. If $\gamma_{V_{k[\sqrt{-a}]}}(\cdot, \cdot)$ is nondegenerate as a bilinear form on $V_{k[\sqrt{-a}]}$, then (12.8.1) and (12.8.2) are nondegenerate as bilinear forms on $V_{k}$.

Let $\alpha_{V_{k}}(\cdot, \cdot)$ be a bilinear form on $V_{k}$ such that

\begin{equation}
\alpha_{V_{k}}(J_{a}(v), w) = \alpha_{V_{k}}(v, J_{a}(w))
\end{equation}

for every $v, w \in V$. Put

\begin{equation}
\alpha_{V_{k[\sqrt{-a}]}}(v, w) = \alpha_{V_{k}}(v, w) + (1/\sqrt{-a}) \alpha_{V_{k}}(J_{a}(v), w)
\end{equation}

for every $v, w \in V$. One can check that this defines a bilinear form on $V_{k[\sqrt{-a}]}$. If $\alpha_{V_{k}}(\cdot, \cdot)$ is symmetric or antisymmetric, then (12.8.3) and (12.8.4) have the same property. Similarly, if $\alpha_{V_{k}}(\cdot, \cdot)$ is nondegenerate, then (12.8.3) and (12.8.4) are nondegenerate.

Suppose for the moment that $\gamma_{V_{k[\sqrt{-a}]}}(\cdot, \cdot)$ is nondegenerate on $V_{k[\sqrt{-a}]}$, so that $\gamma_{V_{k}}(\cdot, \cdot)$ is nondegenerate on $V_{k}$. If $T \in \mathcal{L}_{k[\sqrt{-a}]}(V_{k[\sqrt{-a}]})$, then there is a unique $T^{*}\gamma_{V_{k[\sqrt{-a}]}} \in \mathcal{L}_{k[\sqrt{-a}]}(V_{k[\sqrt{-a}]})$ such that

\begin{equation}
\gamma_{V_{k[\sqrt{-a}]}}(v, w) = \gamma_{V_{k[\sqrt{-a}]}}(v, T^{*}\gamma_{V_{k[\sqrt{-a}]}}(w))
\end{equation}

for every $v, w \in V$. Similarly, if $T \in \mathcal{L}_{k}(V_{k})$, then there is a unique $T^{*}\gamma_{k}$ in $\mathcal{L}_{k}(V_{k})$ such that

\begin{equation}
\gamma_{V_{k}}(v, w) = \gamma_{V_{k}}(v, T^{*}\gamma_{k}(w))
\end{equation}

for every $v, w \in V$. If $T \in \mathcal{L}_{k[\sqrt{-a}]}(V_{k[\sqrt{-a}]})$, then $T \in \mathcal{L}_{k}(V_{k})$, and

\begin{equation}
T^{*}\gamma_{k} = T^{*}\gamma_{k[\sqrt{-a}]}.
\end{equation}

Note that

\begin{equation}
J_{a}^{*}\gamma_{k} = J_{a}^{*}\gamma_{k[\sqrt{-a}]} = J_{a}.
\end{equation}
12.9. CONJUGATE-LINEARITY AND ADJOINTS

Let $\langle \cdot, \cdot \rangle_{\gamma_k[\sqrt{-a}]}$ be an inner product on $V_{k[\sqrt{-a}]}$. If $w \in V$, then there is a unique element $C(w)$ of $V$ such that
\[(12.8.9) \quad \gamma_{\gamma_k[\sqrt{-a}]}(v, w) = \langle v, C(w) \rangle_{\gamma_k[\sqrt{-a}]}
\]
for every $v \in V$, by standard arguments. More precisely, the left side may be considered as a linear functional on $V_{k[\sqrt{-a}]}$ as a function of $v$, which can be represented in terms of the inner product as on the right side. It is easy to see that $C$ is conjugate-linear as a mapping from $V$ into itself, because the left side of (12.8.9) is linear over $k[\sqrt{-a}]$ in $w$, and the inner product $\langle \cdot, \cdot \rangle_{\gamma_k[\sqrt{-a}]}$ is conjugate-linear in the second variable.

Suppose that $\gamma_{\gamma_k[\sqrt{-a}]}(\cdot, \cdot)$ is nondegenerate on $V_{k[\sqrt{-a}]}$ again, which implies that $C$ is a one-to-one mapping from $V$ onto itself. If $T$ is a $k[\sqrt{-a}]$-linear mapping from $V$ into itself, then
\[(12.8.10) \quad (T(v), C(w))_{\gamma_k[\sqrt{-a}]} = \langle v, C(T^* \gamma_k[\sqrt{-a}](w)) \rangle_{\gamma_k[\sqrt{-a}]}
\]
for every $v, w \in V$, by (12.8.5) and (12.8.9). We also have that
\[(12.8.11) \quad (T(v), C(w))_{\gamma_k[\sqrt{-a}]} = \langle v, T^* V_k[\sqrt{-a}](C(w)) \rangle_{\gamma_k[\sqrt{-a}]}
\]
for every $v, w \in V$, where $T^* V_k[\sqrt{-a}]$ is the adjoint of $T$ with respect to the inner product $\langle \cdot, \cdot \rangle_{\gamma_k[\sqrt{-a}]}$ on $V_{k[\sqrt{-a}]}$. It follows that
\[(12.8.12) \quad C \circ T^* V_k[\sqrt{-a}] = T^* V_k[\sqrt{-a}] \circ C.
\]

12.9 Conjugate-linearity and adjoints

Let $k$ be an ordered field, let $a$ be a positive element of $k$, and let $k[\sqrt{-a}]$ be as in Section 12.1. Also let $V = V_{k[\sqrt{-a}]}$ be a finite-dimensional vector space over $k[\sqrt{-a}]$, and let $V_k$ be $V$ considered as a vector space over $k$, as before. Note that the composition of two conjugate-linear mappings from $V$ into itself is linear over $k[\sqrt{-a}]$. Similarly, the composition of a conjugate-linear mapping from $V$ into itself with a $k[\sqrt{-a}]$-linear mapping from $V$ into itself, in either order, is conjugate-linear.

Let $\langle \cdot, \cdot \rangle_{\gamma_k[\sqrt{-a}]}$ be an inner product on $V_{k[\sqrt{-a}]}$, and let $\langle \cdot, \cdot \rangle_{V_k}$ be its real part, as in (12.7.3). Thus $\langle \cdot, \cdot \rangle_{V_k}$ is an inner product on $V_k$ as a vector space over $k$, as before. Let $C$ be a conjugate-linear mapping from $V$ into itself, which is linear over $k$ in particular. The adjoint $C^* V_k$ of $C$ with respect to $\langle \cdot, \cdot \rangle_{V_k}$ is defined as a $k$-linear mapping from $V$ into itself in the usual way. Remember that $J_a$ is the mapping from $V$ into itself defined by multiplication by $\sqrt{-a}$, as in (12.7.7). Conjugate-linearity of $C$ means that
\[(12.9.1) \quad C \circ J_a = -J_a \circ C,
\]
as in Section 12.1. This implies that

\[(12.9.2) \quad C^{*V_k} \circ J_a = - J_a \circ C^{*V_k},\]

because \(J_a^{*V_k} = - J_a\), as in (12.7.10). This shows that \(C^{*V_k}\) is conjugate-linear on \(V\) as well.

Observe that

\[(12.9.3) \quad \langle v, C(w) \rangle_{V_k} = \langle C^{*V_k}(v), w \rangle_{V_k} = \langle w, C^{*V_k}(v) \rangle_{V_k}\]

for every \(v, w \in V\), by definition of the adjoint. In fact, we have that

\[(12.9.4) \quad \langle v, C(w) \rangle_{k[\sqrt{-a}]} = \langle w, C^{*V_k}(v) \rangle_{k[\sqrt{-a}]}\]

for every \(v, w \in V\). More precisely, the real parts of both sides are the same, by (12.9.3). One can get (12.9.4) using this and the fact that both sides are \(k[\sqrt{-a}]\)-linear in \(w\). This gives another way to see the conjugate-linearity of \(C^{*V_k}\) on \(V\) too.

Put

\[(12.9.5) \quad \gamma_{k[\sqrt{-a}]}(v, w) = \langle v, C(w) \rangle_{k[\sqrt{-a}]},\]

for every \(v, w \in V\). This defines a bilinear form on \(V_{k[\sqrt{-a}]}\), because \(C\) is conjugate-linear, and \(\langle \cdot, \cdot \rangle_{k[\sqrt{-a}]}\) is sesquilinear. Similarly, put

\[(12.9.6) \quad \gamma_{V_k}(v, w) = \langle v, C(w) \rangle_{V_k}\]

for every \(v, w \in V\), which is the same as the real part of (12.9.5). Thus

\[(12.9.7) \quad \gamma_{k[\sqrt{-a}]}(v, w) = \langle w, C^{*V_k}(v) \rangle_{k[\sqrt{-a}]},\]

and

\[(12.9.8) \quad \gamma_{V_k}(v, w) = \langle w, C^{*V_k}(v) \rangle_{V_k}\]

for every \(v, w \in V\). It follows that the self-adjointness or anti-self-adjointness of \(C\) with respect to \(\langle \cdot, \cdot \rangle_{V_k}\) is equivalent to the symmetry or anti-symmetry of (12.9.5), (12.9.6), as appropriate.

### 12.10 Antisymmetric forms over \(k[\sqrt{-a}]\)

Let \(k\) be an ordered field, let \(a\) be a positive element of \(k\), and let \(k[\sqrt{-a}]\) be as in Section 12.1. Also let \(V\) be a vector space over \(k[\sqrt{-a}]\) of positive finite dimension, and let \(\gamma(\cdot, \cdot)\) be a nondegenerate antisymmetric bilinear form on \(V\). As in Section 11.13, there is a basis for \(V\) consisting of vectors \(x_1, \ldots, x_n, y_1, \ldots, y_n\) for some positive integer \(n\) such that

\[(12.10.1) \quad \gamma(x_j, x_l) = \gamma(y_j, y_l) = 0\]

for every \(j, l = 1, \ldots, n\),

\[(12.10.2) \quad \gamma(x_j, y_l) = 0\]
when \( j \neq l \), and
\[
\gamma(x_j, y_j) = 1
\]
for every \( j = 1, \ldots, n \). It is easy to define an inner product on \( V \), as a vector space over \( k[\sqrt{-a}] \), for which \( x_1, \ldots, x_n, y_1, \ldots, y_n \) are orthonormal, as in the next paragraph. Using this, we can express \( \gamma(\cdot, \cdot) \) in terms of a conjugate-linear mapping \( C \) from \( V \) into itself, as in the previous sections.

Each \( v \in V \) can be expressed in a unique way as
\[
v = \sum_{j=1}^{n} v_{x_j} x_j + \sum_{j=1}^{n} v_{y_j} y_j,
\]
where \( v_{x_j}, v_{y_j} \in k[\sqrt{-a}] \) for every \( j = 1, \ldots, n \). Similarly, if \( w \in V \), then let \( w_{x_j}, w_{y_j} \in k[\sqrt{-a}] \) be as in (12.10.4). Put
\[
\langle v, w \rangle_V = \sum_{j=1}^{n} v_{x_j} \overline{w_{x_j}} + \sum_{j=1}^{n} v_{y_j} \overline{w_{y_j}}
\]
for every \( v, w \in V \). It is easy to see that this defines an inner product on \( V \), as a vector space over \( k[\sqrt{-a}] \). Note that \( x_1, \ldots, x_n, y_1, \ldots, y_n \) are orthonormal in \( V \) with respect to (12.10.5), by construction.

Let \( C \) be the unique conjugate-linear mapping from \( V \) into itself such that
\[
C(x_j) = -y_j, \quad C(y_j) = x_j
\]
for each \( j = 1, \ldots, n \). Thus, if \( v \in V \) is as in (12.10.4), then
\[
C(v) = -\sum_{j=1}^{n} \overline{v_{y_j}} y_j + \sum_{j=1}^{n} \overline{v_{x_j}} x_j.
\]
One can verify that
\[
\gamma(v, w) = \langle v, C(w) \rangle_V
\]
for every \( v, w \in V \). More precisely, one can first check that this holds when \( v, w \) are among the basis vectors \( x_1, \ldots, x_n, y_1, \ldots, y_n \). This implies that (12.10.8) holds for all \( v, w \in V \), because both sides of (12.10.8) are \( k[\sqrt{-a}] \)-linear in \( v \) and \( w \).

Observe that
\[
C^2 = -I_V,
\]
where \( I_V \) is the identity mapping on \( V \). Let \( V_k \) be \( V \) considered as a vector space over \( k \), as before. Also let \( \langle \cdot, \cdot \rangle_{V_k} \) be the real part of (12.10.5), which defines an inner product on \( V_k \). One can check that \( C \) is anti-self-adjoint as a \( k \)-linear mapping from \( V \) into itself with respect to \( \langle \cdot, \cdot \rangle_{V_k} \). This corresponds to the antisymmetry of (12.10.8) as a bilinear form on \( V \), as in the previous section.
12.11 Symmetric forms over \( k[\sqrt{-a}] \)

Let \( k \) be an ordered field, let \( a \) be a positive element of \( k \), and let \( k[\sqrt{-a}] \) be as in Section 12.1 again. Also let \( V \) be a vector space over \( k[\sqrt{-a}] \) of positive finite dimension \( n \), and let \( e_1, \ldots, e_n \) be a basis for \( V \). Suppose that \( \gamma(\cdot, \cdot) \) is a bilinear form on \( V \) such that
\[
\gamma(e_j, e_l) = 0 \quad \text{when } j \neq l, \text{ and for each } j = 1, \ldots, n, \tag{12.11.1}
\]
and
\[
\gamma(e_j, e_j) \neq 0 \tag{12.11.2}
\]
and
\[
\gamma(e_j, e_j) \in k. \tag{12.11.3}
\]

Note that \( \gamma(\cdot, \cdot) \) is symmetric and nondegenerate on \( V \), because of (12.11.1) and (12.11.2). Conversely, it is well known that for any nondegenerate symmetric bilinear form on \( V \), there is a basis for \( V \) for which (12.11.1) and (12.11.2) hold, as in Sections 11.12 and 11.14. Because of (12.11.3), we can reduce further to the case where
\[
\gamma(e_j, e_j) > 0 \tag{12.11.4}
\]
for every \( j = 1, \ldots, n \), using scalar multiplication by \( \sqrt{-a} \), when needed. If \( k = \mathbb{R} \), so that \( k[\sqrt{-a}] = \mathbb{C} \), then we can reduce directly to the case where
\[
\gamma(e_j, e_j) = 1 \tag{12.11.5}
\]
for every \( j = 1, \ldots, n \), without asking that (12.11.3) hold.

Every \( v \in V \) can be expressed in a unique way as
\[
v = \sum_{j=1}^{n} v_j e_j, \tag{12.11.6}
\]
where \( v_j \in k[\sqrt{-a}] \) for each \( j = 1, \ldots, n \). If \( v, w \in V \), then put
\[
\langle v, w \rangle_V = \sum_{j=1}^{n} |\gamma(e_j, e_j)| v_j \overline{w_j}, \tag{12.11.7}
\]
where \( w_j \in k[\sqrt{-a}] \) corresponds to \( w \) as in (12.11.6). Remember that the absolute value \( |\gamma(e_j, e_j)| \) of \( \gamma(e_j, e_j) \) is defined as an element of \( k \) as in Section 8.13. Of course, this is the same as \( \gamma(e_j, e_j) \) when (12.11.4) holds. It is easy to see that (12.11.7) defines an inner product on \( V \), for which the \( e_j \)'s are pairwise orthogonal.

Let \( C \) be the unique conjugate-linear mapping from \( V \) into itself such that
\[
C(e_j) = \begin{cases} e_j & \text{when } \gamma(e_j, e_j) > 0 \\ -e_j & \text{when } -\gamma(e_j, e_j) > 0 \end{cases}. \tag{12.11.8}
\]
12.12. NONNEGATIVE SELF-ADJOINT OPERATORS

If (12.11.4) holds, then $C(e_j) = e_j$ for every $j = 1, \ldots, n$, so that

(12.11.9) \[ C(v) = \sum_{j=1}^{n} v_j e_j \]

for every $v \in V$ as in (12.11.6). Observe that

(12.11.10) \[ \langle v, C(w) \rangle_V = \sum_{j=1}^{n} \gamma(e_j, e_j) v_j w_j \]

for every $v, w \in V$. Using this, one can check that

(12.11.11) \[ \gamma(v, w) = \langle v, C(w) \rangle_V \]

for every $v, w \in V$. One can also verify that

(12.11.12) \[ C^2 = I_V, \]

where $I_V$ is the identity mapping on $V$.

Let $V_k$ be $V$ considered as a vector space over $k$, as usual. The real part $\langle v, w \rangle_{V_k}$ of (12.11.7) defines an inner product on $V_k$, as before. One can check that $C$ is self-adjoint with respect to $\langle \cdot, \cdot \rangle_{V_k}$, as a $k$-linear mapping from $V$ into itself. This corresponds to the fact that (12.11.11) is symmetric on $V$, as in Section 12.9.

12.12 Nonnegative self-adjoint operators

Let $(V, \langle \cdot, \cdot \rangle_V)$ and $(W, \langle \cdot, \cdot \rangle_W)$ be finite-dimensional inner product spaces, both defined over the real numbers or both defined over the complex numbers, and let $\| \cdot \|_V$, $\| \cdot \|_W$ be the corresponding norms on $V, W$, respectively. Also let $T$ be a linear mapping from $V$ into $W$, and let $T^*$ be the corresponding adjoint mapping from $W$ into $V$. Thus $T^* \circ T$ maps $V$ into itself, and it is easy to see that $T^* \circ T$ is self-adjoint. Observe that

(12.12.1) \[ \langle (T^* \circ T)(v), v \rangle_V = \| T(v) \|_W^2 \]

for every $v \in V$. Of course, one could consider infinite-dimensional Hilbert spaces as well.

Let $A$ be a linear mapping from $V$ into itself that is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_V$, so that

(12.12.2) \[ \langle A(v), w \rangle = \langle v, A(w) \rangle \]

for every $v, w \in V$. In the complex case, (12.12.2) implies that

(12.12.3) \[ \langle A(v), v \rangle = \langle v, A(v) \rangle = \langle A(v), v \rangle \]
for every \( v \in V \), so that \( \langle A(v), v \rangle \in \mathbb{R} \) for every \( v \in V \). In both the real and complex cases, \( A \) is said to be nonnegative on \( V \) if

\[
\langle A(v), v \rangle \geq 0
\]

for every \( v \in V \). Similarly, \( A \) is said to be strictly positive on \( V \) if

\[
\langle A(v), v \rangle > 0
\]

for every \( v \in V \) with \( v \neq 0 \). In particular, this implies that the kernel of \( A \) is trivial. If \( T \) is a linear mapping from \( V \) into \( W \), as before, then \( T^* \circ T \) is automatically nonnegative, by (12.12.1). More precisely, \( T^* \circ T \) is strictly positive exactly when the kernel of \( T \) is trivial.

If \( A \) is any self-adjoint linear mapping from \( V \) into itself, then there is an orthonormal basis for \( V \) consisting of eigenvectors for \( A \), because \( V \) has finite dimension. Note that the eigenvalues of \( A \) are real, even in the complex case. It is easy to see that \( A \) is nonnegative on \( V \) if and only if the eigenvalues of \( A \) are nonnegative real numbers. Similarly, \( A \) is strictly positive on \( V \) if and only if the eigenvalues of \( A \) are positive real numbers. If \( A \) is nonnegative on \( V \) and the kernel of \( A \) is trivial, then it follows that \( A \) is strictly positive on \( V \).

If \( A \) is nonnegative on \( V \), then there is a nonnegative self-adjoint linear mapping \( B \) from \( V \) into itself such that \( B^2 = A \). This can be obtained using a diagonalization for \( A \), as in the preceding paragraph. If \( A \) is strictly positive on \( V \), then \( B \) is strictly positive on \( V \) as well. Note that \( B \) automatically commutes with \( A \) in this situation. More precisely, \( B \) commutes with any linear mapping \( C \) from \( V \) into itself that commutes with \( A \). Indeed, if \( C \) commutes with \( A \), then \( C \) maps the eigenspaces of \( A \) into themselves. On each eigenspace of \( A \), \( B \) is equal to a nonnegative multiple of the identity, by construction.

Let \( B_1 \) be any nonnegative self-adjoint linear mapping from \( V \) into itself such that \( B_1^2 = A \). Thus \( B_1 \) commutes with \( A \), so that \( B_1 \) maps the eigenspaces of \( A \) into themselves. One can check that the restriction of \( B_1 \) to each eigenspace of \( A \) is a nonnegative multiple of the identity. This can be obtained using a diagonalization of \( B_1 \) with respect to an orthonormal basis for each eigenspace of \( A \). It follows that \( B_1 \) is uniquely determined by \( A \) under these conditions.

### 12.13 Polar decompositions

Let us continue with the same notation and hypotheses as in the previous section. Let \( T \) be a linear mapping from \( V \) into \( W \) again, so that \( T^* \circ T \) is a nonnegative self-adjoint linear mapping from \( V \) into itself. It follows that there is a unique nonnegative self-adjoint linear mapping \( R \) from \( V \) into itself such that

\[
R^2 = T^* \circ T,
\]

as in the previous section. If \( u, v \in V \), then

\[
\langle (T^* \circ T)(u), v \rangle_V = \langle T(u), T(v) \rangle_W
\]
12.13. POLAR DECOMPOSITIONS

and
\[ \langle R^2(u), v \rangle_V = \langle R(u), R(v) \rangle_V. \]

This implies that
\[ \langle R(u), R(v) \rangle_V = \langle T(u), T(v) \rangle_W, \]
by (12.13.1). In particular,
\[ \|R(v)\|_V = \|T(v)\|_W \]
for every \( v \in V \), by taking \( u = v \) in (12.13.4). Let us suppose from now on in this section that the kernel of \( T \) is trivial. This implies that the kernel of \( R \) is trivial too, by (12.13.5). Thus \( R \) is invertible as a mapping from \( V \) into itself, because \( V \) has finite dimension.

Put
\[ U = T \circ R^{-1}, \]
which defines a linear mapping from \( V \) into \( W \). If \( v, v' \in V \), then
\[ \langle U(R(v)), U(R(v')) \rangle_V = \langle T(v), T(v') \rangle_W = \langle R(v), R(v') \rangle_V, \]
using the definition of \( U \) in the first step, and (12.13.4) in the second step. This means that
\[ \langle U(v), U(v') \rangle_W = \langle v, v' \rangle_V \]
for every \( v, v' \in V \), because \( R \) is invertible on \( V \). It follows that
\[ \|U(v)\|_W = \|v\|_V \]
for every \( v \in V \), by taking \( v' = v \) in (12.13.8). Let us suppose from now on in this section that \( V = W \). This implies that \( T \) maps \( V \) onto itself, because \( V \) has finite dimension. Thus \( U \) is a one-to-one mapping from \( V \) onto itself. More precisely,
\[ U^{-1} = U^*, \]
because of (12.13.8), as in Section 3.8.

Suppose that \( T \) is normal on \( V \), in the sense that \( T \) commutes with \( T^* \). Of course, this implies that \( T \) commutes with \( T^* \circ T \). It follows that \( R \) commutes with \( T \), as in the previous section. This means that \( U \) commutes with \( R \) and \( T \), by the definition of \( U \).

Suppose for the moment that \( T \) is self-adjoint on \( V \), which implies that \( T \) is normal in particular. In this case, we get that
\[ U^2 = T^2 \circ R^{-2} = I, \]
the identity operator on \( V \). Equivalently, one can check that \( U \) is self-adjoint in this situation.

Similarly, if \( T \) is anti-self-adjoint on \( V \), then \( T \) is normal, and \( T^* \circ T = -T^2 \). This implies that
\[ U^2 = T^2 \circ R^{-2} = -I. \]
Alternatively, one can verify directly that \( U \) is anti-self-adjoint in this case.
12.14 Bilinear forms and inner products

Let \((V, \langle \cdot, \cdot \rangle)\) be a finite-dimensional inner product space over the real numbers. If \(B\) is a linear mapping from \(V\) into itself, then

\[
\beta(v, w) = \langle B(v), w \rangle
\]  

(12.14.1)

defines a bilinear form on \(V\). Every bilinear form on \(V\) corresponds to a unique linear mapping on \(V\) in this way. Note that \(\beta\) is symmetric or antisymmetric on \(V\) exactly when \(B\) is self-adjoint or anti-self-adjoint on \(V\) with respect to \(\langle \cdot, \cdot \rangle\), as appropriate. Let us suppose from now on that \(B\) is invertible on \(V\), so that \(\beta\) is nondegenerate on \(V\).

Of course, the adjoint \(B^*\) of \(B\) with respect to \(\langle \cdot, \cdot \rangle\) is invertible on \(V\) too, because \(B\) is invertible. Note that \(B \circ B^*\) is a strictly positive self-adjoint linear mapping from \(V\) into itself, which corresponds to taking \(T = B^*\) in the previous section. As before, there is a unique nonnegative self-adjoint linear mapping \(A\) from \(V\) into itself such that

\[
A^2 = B \circ B^*.
\]  

(12.14.2)

More precisely, \(A\) is invertible on \(V\), so that \(A\) is strictly positive on \(V\). Put

\[
B_0 = A^{-1} \circ B,
\]  

(12.14.3)

which is invertible on \(V\) as well. By construction,

\[
B_0 \circ B_0^* = A^{-1} \circ B \circ B^* \circ A^{-1} = I,
\]  

(12.14.4)

so that

\[
B_0^{-1} = B_0^*.
\]  

(12.14.5)

If we take \(T = B^*\) in the previous section, then \(A\) corresponds to \(R\), and \(B_0\) corresponds to \(U^*\).

If \(B\) is normal on \(V\), then \(A\) commutes with \(B\), and hence \(B_0\) commutes with \(A\) and \(B\), as before. If \(B\) is self-adjoint on \(V\), and normal in particular, then

\[
B_0^2 = A^{-2} \circ B^2 = I.
\]  

(12.14.6)

Equivalently, one can check that \(B_0\) is self-adjoint on \(V\). Similarly, if \(B\) is anti-self-adjoint on \(V\), then \(B\) is normal, and

\[
B_0^2 = A^{-2} \circ B^2 = -I.
\]  

(12.14.7)

Alternatively, one can verify that \(B_0\) is anti-self-adjoint in this case.

Clearly

\[
B = A \circ B_0,
\]  

(12.14.8)

by (12.14.3). Put

\[
\langle v, w \rangle_A = \langle A(v), w \rangle
\]  

(12.14.9)
12.15. COMPLEX BILINEAR FORMS

for every \(v, w \in V\). This defines an inner product on \(V\), because \(A\) is strictly positive and self-adjoint on \(V\). Observe that

\[
\beta(v, w) = \langle B_0(v), w \rangle_A
\]

for every \(v, w \in V\). If \(B\) is normal with respect to \(\langle \cdot, \cdot \rangle\) on \(V\), then

\[
\langle B_0(v), w \rangle_A = \langle A(B_0(v)), w \rangle = \langle B_0(A(v)), w \rangle = \langle A(v), B_0(w) \rangle = \langle v, B_0(w) \rangle_A
\]

for every \(v, w \in V\), because \(B_0\) commutes with \(A\).

Now let \((V, \langle \cdot, \cdot \rangle)\) be a finite-dimensional inner product space over the complex numbers. If \(C\) is a conjugate-linear mapping from \(V\) into itself, then (12.14.1) defines a sesquilinear form \(\beta\) on \(V\), and every sesquilinear form on \(V\) corresponds to a unique linear mapping from \(V\) into itself in this way. One can check that \(\beta\) is Hermitian-symmetric on \(V\) exactly when \(B\) is self-adjoint on \(V\) with respect to \(\langle \cdot, \cdot \rangle\). Suppose from now on in this section that \(B\) is invertible on \(V\), so that \(\beta\) is nondegenerate on \(V\). This implies that the adjoint \(B^*\) of \(B\) with respect to \(\langle \cdot, \cdot \rangle\) is invertible on \(V\) as well.

As before, \(B \circ B^*\) is a strictly positive self-adjoint linear mapping from \(V\) into itself. Hence there is a unique nonnegative self-adjoint linear mapping \(A\) from \(V\) into itself that satisfies (12.14.2). In fact, \(A\) is invertible on \(V\), and strictly positive. Let \(B_0\) be as in (12.14.3), which is invertible on \(V\), with inverse equal to \(B_0^*\), as in (12.14.5). If \(B\) is normal on \(V\), then \(A\) commutes with \(B\), and \(B_0\) commutes with \(A\) and \(B\). If \(B\) is self-adjoint on \(V\), then \(B\) is normal, and

\[
B_0^2 = I.
\]

Alternatively, one can check directly that \(B_0\) is self-adjoint in this case.

Let \(\langle v, w \rangle_A\) be defined for \(v, w \in V\) as in (12.14.9), which defines an inner product on \(V\), because \(A\) is strictly positive and self-adjoint on \(V\). Thus (12.14.10) holds for every \(v, w \in V\) again, by construction. If \(B\) is normal on \(V\) with respect to \(\langle \cdot, \cdot \rangle\), then the adjoint of \(B_0\) with respect to \(\langle \cdot, \cdot \rangle_A\) is the same as the adjoint with respect to \(\langle \cdot, \cdot \rangle\), as in (12.14.11).

12.15 Complex bilinear forms

Let \((V, \langle \cdot, \cdot \rangle)\) be a finite-dimensional inner product space over the complex numbers again. If \(C\) is a conjugate-linear mapping from \(V\) into itself, then

\[
\gamma(v, w) = \langle v, C(w) \rangle
\]

is a bilinear form on \(V\), as before. Every bilinear form on \(V\) corresponds to a conjugate-linear mapping \(C\) on \(V\) in this way, as in Section 12.8.

Let \(V_R\) be \(V\) considered as a vector space over the real numbers, and let \(\langle \cdot, \cdot \rangle_{V_R}\) be the real part of \(\langle \cdot, \cdot \rangle\), which is an inner product on \(V_R\). Also let \(C\) be a conjugate-linear mapping from \(V\) into itself, and let \(C^{*V_R}\) be the adjoint
of $C$, as a real-linear mapping from $V$ into itself, with respect to $\langle \cdot, \cdot \rangle_{V_R}$. Thus $C^{*}\cdot V_R$ is conjugate-linear on $V$ too, as in Section 12.9. It follows that $C \circ C^{*}\cdot V_R$ is a complex-linear mapping from $V$ into itself, as before.

Note that $C \circ C^{*}\cdot V_R$ is nonnegative and self-adjoint with respect to $\langle \cdot, \cdot \rangle_{V_R}$. This implies that there is a unique real-linear mapping $A$ from $V$ into itself that is nonnegative and self-adjoint with respect $\langle \cdot, \cdot \rangle_{V_R}$ and satisfies

$$A^2 = C \circ C^{*}\cdot V_R.$$  

(12.15.2)

Remember that $A$ commutes with any real-linear mapping from $V$ into itself that commutes with $C \circ C^{*}\cdot V_R$. In particular, $A$ commutes with the mapping $J$ from $V$ into itself that corresponds to multiplication by $i$, because $C \circ C^{*}\cdot V_R$ is complex linear. This means that $A$ is complex linear on $V$ as well, and one can check that $A$ is self-adjoint and nonnegative with respect to $\langle \cdot, \cdot \rangle_{V_R}$. Alternatively, $C \circ C^{*}\cdot V_R$ is self-adjoint as a complex-linear mapping from $V$ into itself with respect to $\langle \cdot, \cdot \rangle_{V_R}$, because $C \circ C^{*}\cdot V_R$ is complex linear and self-adjoint as a real-linear mapping from $V$ into itself with respect to $\langle \cdot, \cdot \rangle_{V_R}$, as in Section 12.7. Similarly, it is easy to see that $C \circ C^{*}\cdot V_R$ is nonnegative as a self-adjoint complex-linear mapping from $V$ into itself with respect to $\langle \cdot, \cdot \rangle_{V_R}$. Thus one can take $A$ to be the unique complex-linear mapping from $V$ into itself that is nonnegative and self-adjoint with respect to $\langle \cdot, \cdot \rangle_{V_R}$ and satisfies (12.15.2).

Suppose from now on in this section that $C$ is invertible on $V$, so that (12.15.1) is nondegenerate as a bilinear form on $V$. This implies that $C^{*}\cdot V_R$ is invertible on $V$ too, and that $C \circ C^{*}\cdot V_R$ is strictly positive on $V$ with respect to $\langle \cdot, \cdot \rangle_{V_R}$, and hence with respect to $\langle \cdot, \cdot \rangle_{V_R}$. It follows that $A$ is invertible on $V$, and strictly positive with respect to $\langle \cdot, \cdot \rangle_{V_R}, \langle \cdot, \cdot \rangle_{V_R}$. Thus

$$\langle v, w \rangle_{V_R, A} = \langle A(v), w \rangle_{V_R}$$  

(12.15.3)

defines an inner product on $V_R$, and

$$\langle v, w \rangle_{A} = \langle A(v), w \rangle$$  

(12.15.4)

defines an inner product on $V$. Of course,

$$\langle v, w \rangle_{V_R, A} = \text{Re} \langle v, w \rangle_{A}$$  

(12.15.5)

for every $v, w \in V$, by construction.

Put

$$C_0 = A^{-1} \circ C,$$  

(12.15.6)

which is invertible as a real-linear mapping on $V$. In fact, $C_0$ is conjugate-linear on $V$, because $A$ is complex-linear and $C$ is conjugate-linear. As before, $C_0 \circ C_0^{*}\cdot V_R = I$, so that

$$C_0^{-1} = C_0^{*}\cdot V_R.$$  

(12.15.7)
12.15. COMPLEX BILINEAR FORMS

Of course, \( C = A \circ C_0 \), so that

\[
\gamma(v, w) = \langle v, C(w) \rangle = \langle v, A(C_0(w)) \rangle = \langle A(v), C_0(w) \rangle = \langle v, C_0(w) \rangle_A
\]

for every \( v, w \in V \).

Suppose for the moment that \( C \) is normal with respect to \( \langle \cdot, \cdot \rangle_{V_R} \), so that \( C \) commutes with \( C^{*, V_R} \). This implies that \( C \) commutes with \( C \circ C^{*, V_R} \), and hence that \( A \) commutes with \( C \). It follows that \( C_0 \) commutes with \( A \) and \( C \) in this situation. Thus

\[
\langle C_0(v), w \rangle_{V_{R,A}} = \langle A(v), C_0(w) \rangle_{V_R} = \langle C_0(A(v)), w \rangle_{V_R} = \langle v, C^{*, V_R}(w) \rangle_{V_{R,A}}
\]

for every \( v, w \in V \). This means that the adjoint of \( C_0 \) with respect to \( \langle \cdot, \cdot \rangle_{V_{R,A}} \) is the same as \( C_0^{*, V_R} \) in this case.

If \( \gamma(\cdot, \cdot) \) is symmetric or antisymmetric on \( V \), then

\[
\gamma_R(v, w) = \text{Re} \gamma(v, w) = \langle v, C(w) \rangle_{V_R}
\]

has the analogous property as a bilinear form on \( V_R \). This means that \( C \) is self-adjoint or anti-self-adjoint with respect to \( \langle \cdot, \cdot \rangle_{V_R} \). In both cases, \( C \) is normal with respect to \( \langle \cdot, \cdot \rangle_{V_R} \). If \( C \) is self-adjoint with respect to \( \langle \cdot, \cdot \rangle_{V_R} \), then \( C_0 \) is self-adjoint with respect to \( \langle \cdot, \cdot \rangle_{V_R} \), and

\[
C_0^2 = I,
\]

as in the previous section. Similarly, if \( C \) is anti-self-adjoint with respect to \( \langle \cdot, \cdot \rangle_{V_R} \), then \( C_0 \) is anti-self-adjoint with respect to \( \langle \cdot, \cdot \rangle_{V_R} \), and

\[
C_0^2 = -I.
\]
Chapter 13

Semisimplicity

13.1 Complete reducibility

Let $k$ be a field, and let $A$ be a Lie algebra over $k$ with finite dimension as a vector space over $k$. Also let $V$ be a finite-dimensional vector space over $k$, and let $\rho$ be a representation of $A$ on $V$. If $V$ corresponds to the direct sum of irreducible representations of $A$, then $\rho$ is said to be completely reducible on $V$. Equivalently, $V$ is said to be semisimple as a module over $A$ in this case.

Let us suppose from now on in this section that $k$ has characteristic 0, and that $A$ is semisimple as a Lie algebra over $k$. This implies that $[A, A] = A$, as in the previous section. It follows that any representation of $A$ on a one-dimensional vector space over $k$ is trivial, as in the lemma on p28 of [13], and remarked on p47 of [24].

A famous theorem going back to Weyl states that every finite-dimensional representation of $A$ is completely reducible, as on p28 of [13] and p46 of [24]. In order to prove this, we shall begin with the following splitting principle. Weyl’s theorem will be obtained from this in Section 13.6.

Let $V$ be a finite-dimensional module over $A$, and suppose that $W$ is a submodule of $V$, as a module over $A$, such that $W$ has codimension one in $V$ as a vector space over $k$. This means that the quotient space $V/W$ has dimension one as a vector space over $k$, so that the induced action of $A$ on $V/W$ is trivial, as before. Equivalently, the action of $A$ on $V$ actually maps $V$ into $W$. Under these conditions, we would like to show that $V$ corresponds to the direct sum of $W$ and a one-dimensional submodule of $V$, as a module over $A$.

The case where $W$ is irreducible as a module over $A$ will be discussed in Section 13.5. This will use a suitable Casimir element, which will be discussed in the next two sections.

To reduce to the case where $W$ is irreducible, we use induction on the dimension of $W$. Suppose that $W$ is not irreducible as a module over $A$, so that there is a proper nonzero submodule $Z$ of $W$, as a module over $A$. Under these conditions, $V/Z$ is an module over $A$ too, $W/Z$ is a submodule of $V/Z$, as a
module over \( A \), and \( W/Z \) has codimension one in \( V/Z \), as a vector space over \( k \). The dimension of \( W/Z \) is strictly less than the dimension of \( W \), as vector spaces over \( k \). Using induction, we get that \( V/Z \) corresponds to the direct sum of \( W/Z \) and a one-dimensional submodule of \( V/Z \), as a module over \( A \).

This one-dimensional submodule of \( V/Z \) can be expressed as \( U/Z \), where \( U \) is a linear subspace of \( V \) that contains \( Z \) as a codimension-one subspace, and \( U \) is a submodule of \( V \), as a module over \( A \). Note that the dimension of \( Z \) is strictly less than the dimension of \( W \), as vector spaces over \( k \). Using the induction hypothesis again, we get that \( U \) corresponds to the direct sum of \( Z \) and a one-dimensional submodule of \( U \), as a module over \( A \). One can check that \( V \) corresponds to the direct sum of \( W \) and this one-dimensional submodule of \( U \), as desired.

### 13.2 Casimir elements

Let \( k \) be a field, and let \((A, [\cdot, \cdot]_A)\) be a Lie algebra over \( k \). Suppose that \( A \) has positive finite dimension \( n \) as a vector space over \( k \), and let \( \beta \) be a nondegenerate symmetric bilinear form on \( A \). Suppose also that \( \beta \) is associative on \( A \), or equivalently that \( \beta \) is invariant with respect to the adjoint representation on \( A \), so that

\[
\beta([x, w]_A, y) = \beta(x, [w, y]_A)
\]

for every \( w, x, y \in A \), as in Sections 6.10 and 7.11. Let \( u_1, \ldots, u_n \) be a basis for \( A \), as a vector space over \( k \). Under these conditions, there is a basis \( w_1, \ldots, w_n \) for \( A \) such that

\[
\beta(u_j, u_l) = \delta_{j,l}
\]

for every \( j, l = 1, \ldots, n \). Here \( \delta_{j,l} \in k \) is equal to 1 when \( j = l \), and to 0 when \( j \neq l \), as usual. This is the dual basis for \( A \) with respect to \( \beta \).

Let \( x \in A \) be given, and let \((a_{j,l})\) and \((b_{j,l})\) be the \( n \times n \) matrices with entries in \( k \) such that

\[
[x, u_j]_A = \sum_{l=1}^{n} a_{j,l} u_l
\]

and

\[
[x, w_j]_A = \sum_{l=1}^{n} b_{j,l} w_l
\]

for every \( j = 1, \ldots, n \). Observe that

\[
\beta([x, u_j]_A, w_r) = \sum_{l=1}^{n} a_{j,l} \beta(u_l, w_r) = a_{j,r}
\]

for every \( j, r = 1, \ldots, n \), and similarly

\[
\beta(u_j, [x, w_r]_A) = \sum_{l=1}^{n} b_{r,l} \beta(u_j, w_l) = b_{r,j}
\]
for every $j, r = 1, \ldots, n$. Using (13.2.1), we get that
(13.2.7) $a_{j,r} = -b_{r,j}$
for every $j, r = 1, \ldots, n$.

Let $V$ be a vector space over $k$, and let $\rho$ be a representation of $A$ on $V$. Put
(13.2.8) $c_\rho(\beta) = \sum_{j=1}^{n} \rho_{u_j} \circ \rho_{w_j}$,
which defines a linear mapping from $V$ into itself. This is the Casimir element
of the space $\mathcal{L}(V)$ of linear mappings from $V$ into itself associated to $\beta$ and $\rho$.
If $x \in A$ is as in the preceding paragraph, then
(13.2.9) $[\rho_x, c_\rho(\beta)] = \sum_{j=1}^{n} [\rho_x, \rho_{u_j} \circ \rho_{w_j}]$,
using the commutator bracket on $\mathcal{L}(V)$ corresponding to composition of linear
mappings on $V$. It follows that
(13.2.10) $[\rho_x, c_\rho(\beta)] = \sum_{j=1}^{n} [\rho_x, \rho_{u_j}] \circ \rho_{w_j} + \sum_{j=1}^{n} \rho_{u_j} \circ [\rho_x, \rho_{w_j}]$,
as in Section 2.5. Because $\rho$ is a Lie algebra representation, we get that
(13.2.11) $[\rho_x, c_\rho(\beta)] = \sum_{j=1}^{n} a_{j,l} \rho_{u_j} \circ \rho_{w_l} + \sum_{j=1}^{n} b_{j,l} \rho_{u_j} \circ \rho_{w_l} = 0$,
using (13.2.3) and (13.2.4) in the second step, and (13.2.7) in the third step.
This corresponds to part of the discussion on p27 of [13].

Let $u'_1, \ldots, u'_n$ and $w'_1, \ldots, w'_n$ be bases for $A$ as a vector space over $k$. We
can express these bases in terms of the $u_j$’s and $w_l$’s, so that
(13.2.12) $u'_h = \sum_{j=1}^{n} \mu_{h,j} u_j$
for every $h = 1, \ldots, n$ and
(13.2.13) $w'_r = \sum_{l=1}^{n} \nu_{r,l} w_l$
for every $r = 1, \ldots, n$, where $\mu = (\mu_{h,j})$ and $\nu = (\nu_{r,l})$ are invertible $n \times n$
matrices with entries in $k$. Thus
(13.2.14) $\beta(u'_h, w'_r) = \sum_{j=1}^{n} \sum_{l=1}^{n} \mu_{h,j} \nu_{r,l} \delta_{h,l} = \sum_{j=1}^{n} \mu_{h,j} \nu_{r,j}$
13.3. ANOTHER PERSPECTIVE

for every \( h, r = 1, \ldots, n \), using (13.2.2) in the second step. It follows that

\[
\beta(u'_h, w'_r) = \delta_{h,r}
\]

for every \( h, r = 1, \ldots, n \) if and only if

\[
\sum_{j=1}^{n} \mu_{h,j} \nu_{r,j} = \delta_{h,r}
\]

for every \( h, r = 1, \ldots, n \). Of course, (13.2.16) is the same as saying that \( \mu \) times the transpose of \( \nu \) is the identity matrix. This is equivalent to the condition that the transpose of \( \nu \) times \( \mu \) be the identity matrix, which means that

\[
\sum_{h=1}^{n} \nu_{h,l} \mu_{h,j} = \delta_{l,j}
\]

for every \( j, l = 1, \ldots, n \). If \( \rho \) is as in the previous paragraph, then

\[
\sum_{h=1}^{n} \rho_{u'_h} \circ \rho_{w'_h} = \sum_{h=1}^{n} \sum_{j=1}^{n} \sum_{l=1}^{n} \mu_{h,j} \nu_{h,l} \rho_{u_j} \circ \rho_{w_l}.
\]

If (13.2.17) holds, then we get that

\[
\sum_{h=1}^{n} \rho_{u'_h} \circ \rho_{w'_h} = \sum_{j=1}^{n} \sum_{l=1}^{n} \delta_{l,j} \rho_{u_j} \circ \rho_{w_l} = \sum_{j=1}^{n} \rho_{u_j} \circ \rho_{w_j},
\]

which is the same as (13.2.8).

The hypothesis that \( \beta \) be symmetric on \( A \) does not seem to have been used so far, but it does give some additional properties. If \( \beta \) is symmetric on \( A \), then the conditions on \( u_1, \ldots, u_n \) and \( w_1, \ldots, w_n \) are symmetric in these two bases for \( A \).

13.3 Another perspective

Let us continue with the same notation and hypotheses as in the previous section. Let \( A' \) be the space of linear functionals on \( A \), as a vector space over \( k \). Remember that the tensor products \( A \otimes A \) and \( A \otimes A' \) can be defined as vector spaces over \( k \), as in Section 7.12. The product of an element of \( A \) and a linear functional on \( A \) defines a linear mapping from \( A \) into itself. This defines a bilinear mapping from \( A \times A' \) into the space \( \mathcal{L}(A) \) of linear mappings from \( A \) into itself, as a vector space over \( k \). This leads to a linear mapping from \( A \otimes A' \) into \( \mathcal{L}(A) \), which is an isomorphism in this case, because \( A \) has finite dimension as a vector space over \( k \). If \( z \in A \), then

\[
\beta_z(x) = \beta(x, z)
\]
defines a linear functional on $A$, and $z \mapsto \beta_z$ is a linear mapping from $A$ into $A'$. More precisely, $z \mapsto \beta_z$ is a one-to-one linear mapping from $A$ onto $A'$, because $\beta$ is nondegenerate on $A$. This leads to a one-to-one linear mapping from $A \otimes A$ onto $A \otimes A'$.

The condition (13.2.2) on the bases $u_1, \ldots, u_n$ and $w_1, \ldots, w_n$ for $A$ is the same as saying that

$$\sum_{j=1}^{n} u_j \beta(x, w_j) = x$$

for every $x \in A$. Consider

$$\sum_{j=1}^{n} u_j \otimes w_j,$$

as an element of $A \otimes A$. This corresponds to the identity mapping on $A$ under the isomorphisms mentioned in the preceding paragraph, because of (13.3.2). Remember that $A$ may be considered as a module over itself, as a Lie algebra over $k$, using the adjoint representation. It follows that $A', A$, and the related tensor products may be considered as modules over $A$ too, in the usual way. The invariance condition on $\beta$ implies that $z \mapsto \beta_z$ is a homomorphism from $A$ into $A'$, as modules over $A$. One can use this to get that (13.3.3) is invariant under the corresponding representation of $A$ on $A \otimes A$, because the identity mapping on $A$ is invariant under the representation of $A$ on $L(A)$. This can also be verified using (13.2.3), (13.2.4), and (13.2.7), as before.

The image of (13.3.3) in the universal enveloping algebra of $A$ is called the Casimir element associated to $A$, as on p46 of [24]. One can get (13.2.8) from (13.3.3) using the action of the universal enveloping algebra on $V$ associated to the representation $\rho$, as in [24]. This amounts to using $\rho$ to get a bilinear mapping from $A \times A$ into $L(V)$, and thus a linear mapping from $A \otimes A$ into $L(V)$. Related matters are discussed on p118-9 of [13].

Let $u'_1, \ldots, u'_n$ and $w'_1, \ldots, w'_n$ be bases for $A$ again, which can be expressed in terms of the $u_j$‘s and $w_l$‘s as in (13.2.12) and (13.2.13), respectively. Thus

$$\sum_{h=1}^{n} u'_h \otimes w'_h = \sum_{h=1}^{n} \sum_{j=1}^{n} \sum_{l=1}^{n} u_{h,j} v_{h,l} u_j \otimes w_l.$$ 

If (13.2.17) holds, then it follows that

$$\sum_{h=1}^{n} u'_h \otimes w'_h = \sum_{j=1}^{n} \sum_{l=1}^{n} \delta_{l,j} u_j \otimes w_l = \sum_{j=1}^{n} u_j \otimes w_j.$$ 

This can also be obtained from the fact that (13.3.3) corresponds to the identity mapping on $A$ under the isomorphisms mentioned earlier.

If $\beta$ is symmetric on $A$, then (13.3.3) is symmetric as an element of $A \otimes A$, as in Section 7.15. This follows from (13.3.5), because (13.2.2) is symmetric in $u_1, \ldots, u_n$ and $w_1, \ldots, w_n$ in this case.
13.4 A more particular situation

Let \( k \) be a field of characteristic 0, and let \((A, [\cdot, \cdot]_A)\) be a Lie algebra over \( k \) with positive finite dimension \( n \) as a vector space over \( k \). Also let \( V \) be a finite-dimensional vector space over \( k \), and let \( \rho \) be a representation of \( A \) on \( V \). Put
\[
\beta_\rho(x, y) = \text{tr}_V(\rho_x \circ \rho_y)
\]
for every \( x, y \in A \), which defines a symmetric bilinear form on \( A \). Remember that (13.4.1) satisfies the associativity or invariance condition (13.2.1), as in Section 7.9. Suppose from now on in this section that \( A \) is semisimple as a Lie algebra over \( k \), and that \( \rho \) is injective as a Lie algebra homomorphism from \( A \) into \( gl(V) \).

Remember that the radical of (13.4.1) in \( A \) is defined by
\[
A^{\beta_\rho} = \{ x \in A : \beta_\rho(x, y) = 0 \text{ for every } y \in A \},
\]
as in Section 7.11. This is an ideal in \( A \), because (13.4.1) is associative on \( A \), as before. The image
\[
\{ \rho_x : x \in A^{\beta_\rho} \}
\]
of \( A^{\beta_\rho} \) under \( \rho \) is a Lie subalgebra of \( gl(V) \). Using Cartan’s criterion, we get that (13.4.3) is solvable as a Lie algebra over \( k \). This implies that \( A^{\beta_\rho} = \{0\} \), because \( A \) is semisimple. This shows that (13.4.1) is nondegenerate on \( A \) under these conditions.

Let \( u_1, \ldots, u_n \) be a basis for \( A \), as a vector space over \( k \). As in Section 13.2, there is a basis \( w_1, \ldots, w_n \) for \( A \) that satisfies (13.2.2), with \( \beta = \beta_\rho \). This leads to a Casimir element \( c_\rho = c_\rho(\beta_\rho) \) of \( L(V) \) associated to \( \rho \) as in (13.2.8). Observe that
\[
\text{tr}_V c_\rho = \sum_{j=1}^n \text{tr}_V(\rho_{u_j} \circ \rho_{w_j}) = \sum_{j=1}^n \beta_{\rho}(u_j, w_j) = n.
\]
In particular, this means that \( c_\rho \neq 0 \), because we are assuming for convenience that \( A \neq \{0\} \).

Remember that \( c_\rho \) commutes with the action of \( \rho \) on \( V \), as in (13.2.11). If \( \rho \) is irreducible on \( V \), then it follows that \( c_\rho \) is a one-to-one mapping from \( V \) onto itself, by Schur’s lemma, as in Section 6.14. Otherwise, if \( W \) is a linear subspace of \( V \) that is invariant under the action of \( \rho \), then \( c_\rho \) maps \( W \) into itself as well. This corresponds to parts of the discussions on p27 of [13] and p46 of [24].

13.5 A splitting theorem

Let \( k \) be a field of characteristic 0, and let \( A \) be a finite-dimensional semisimple Lie algebra over \( k \). Also let \( V \) be a finite-dimensional vector space over \( k \), and let \( \rho^V \) be a representation of \( A \) on \( V \). Suppose that \( W \) is a codimension-one linear subspace of \( V \), and that the action of \( \rho^V \) on \( V \) maps \( W \) into itself. We would
like to show that $V$ is the direct sum of $W$ and a one-dimensional subspace that is mapped to itself by $\rho^V$, as in Section 13.1. Remember that the induced action of $\rho^V$ on $V/W$ is trivial, as before.

Let $\rho^W$ be the representation of $A$ on $W$ obtained by restricting $\rho^V$ from $V$ to $W$. In this section, we consider the case where $W$ is irreducible with respect to $\rho^W$. This corresponds to arguments on p28-9 of [13], and p47 of [24]. These arguments focus on $V$ and $W$, respectively, as we shall see. Note that the roles of $V$ and $W$ are exchanged in the notation used on p47 of [24].

If $x, y \in A$, then put

$$\beta_{\rho^V}(x, y) = \text{tr}_V(\rho^V_x \circ \rho^V_y) \quad (13.5.1)$$

and

$$\beta_{\rho^W}(x, y) = \text{tr}_W(\rho^W_x \circ \rho^W_y), \quad (13.5.2)$$

as in (13.4.1). In this situation, (13.5.1) and (13.5.2) are the same, because the action induced on $V/W$ by $\rho^V$ is trivial. This also uses the remarks in Section 7.10.

Let us begin with the argument in [13]. Without loss of generality, we may suppose that $\rho^V$ is injective as a Lie algebra homomorphism from $A$ into $\mathfrak{gl}(V)$. Otherwise, we can replace $A$ with its quotient by the kernel of $\rho^V$. This quotient of $A$ is semisimple as a Lie algebra over $k$ too, as in Section 10.14. We may suppose that $A \neq \{0\}$ as well, since otherwise the problem is very easy.

Let $c_{\rho^V}$ be the Casimir element of $\mathcal{L}(V)$ associated to $\rho^V$ as in the previous section. Thus $c_{\rho^V}$ maps $W$ into itself, because $W$ is invariant under $\rho^V$, by hypothesis. The kernel of $c_{\rho^V}$ in $V$ is invariant under the action of $\rho^V$, because $c_{\rho^V}$ commutes with the action of $\rho^V$ on $V$, as in (13.2.11).

Because $c_{\rho^V}(W) \subseteq W$, $c_{\rho^V}$ induces a linear mapping from $V/W$ into itself. This induced mapping is equal to 0, because of the corresponding statement for $\rho^V$, and the definition of $c_{\rho^V}$. If $c_{\rho^V} \equiv 0$ on $W$, then it follows that $\text{tr}_V c_{\rho^V} = 0$, as in Section 7.10. This contradicts (13.4.4), because $A \neq \{0\}$.

Thus $c_{\rho^V} \neq 0$ on $W$, so that the restriction of $c_{\rho^V}$ to $W$ is a one-to-one mapping onto $W$, by Schur’s lemma, as in Section 6.14. However, $c_{\rho^V}$ is not invertible on $V$, because the induced mapping on $V/W$ is equal to 0. Hence the kernel of $c_{\rho^V}$ is a one-dimensional linear subspace of $V$ whose intersection with $W$ is trivial. This gives an invariant complement of $W$ in $V$, as desired.

Now let us consider the argument in [24]. Let $A_0$ be the kernel of $\rho^W$, as a Lie algebra homomorphism from $A$ into $\mathfrak{gl}(W)$. Thus $A_0$ is an ideal in $A$. If $x \in A_0$, then $\rho^V_x = \rho^W_x = 0$ on $W$, and $\rho^V_x(V) \subseteq W$, because the mapping on $V/W$ induced by $\rho^V_x$ is equal to 0. If $y \in [A_0, A_0]$, then it follows that $\rho^V_y = 0$ on $V$. Remember that $A_0$ is semisimple as a Lie algebra over $k$, because $A$ is semisimple, as in Section 10.14. This implies that $A_0 = [A_0, A_0]$, as in Section 10.15. It follows that $\rho^W_y = 0$ on $V$ for every $y \in A_0$. Note that $A/A_0$ is semisimple as a Lie algebra over $k$, as in Section 10.14 again. This permits us to reduce to the case where $\rho^W$ is injective as a Lie algebra homomorphism from $A$ into $\mathfrak{gl}(W)$, since otherwise we could replace $A$ with $A/A_0$. 


13.6. WEYL’S THEOREM

We may suppose that $A \neq \{0\}$ too, as before. Because $\rho^W$ is injective as a Lie algebra homomorphism from $A$ into $gl(W)$, (13.5.2) is nondegenerate on $A$, as in the previous section. Let $c_{\rho^V}(\beta_{\rho^W})$ be the Casimir element of $\mathcal{L}(V)$ that corresponds to $\rho^V$ and (13.5.2), as in Section 13.2. Note that $c_{\rho^V}(\beta_{\rho^W})$ maps $V$ into $W$, because of the corresponding property of $\rho^V$. The restriction of $c_{\rho^V}(\beta_{\rho^W})$ to $W$ is the same as the Casimir element $c_{\rho^W} = c_{\rho^V}(\beta_{\rho^W})$ of $\mathcal{L}(W)$ that corresponds to $\rho^W$ and (13.5.2), by the definitions of $c_{\rho^V}(\beta_{\rho^W})$ and $\rho^W$. The trace of $c_{\rho^W}$ on $W$ is equal to the dimension of $A$, as in (13.4.4). Thus $c_{\rho^W} \neq 0$, because $A \neq \{0\}$. This implies that $c_{\rho^W}$ is invertible on $W$, by Schur’s lemma, because $\rho^W$ is irreducible on $W$. It follows that the kernel of $c_{\rho^V}(\beta_{\rho^W})$ is a one-dimensional subspace of $V$ complementary to $W$, as before.

More precisely, the argument in [24] is formulated in terms of the Casimir element of the universal enveloping algebra of $A$ associated to (13.5.2), as in Section 13.2. This is used again in the remarks following the proof, with the roles of $V$ and $W$ interchanged in the notation of [24], as before.

13.6 Weyl’s theorem

Let $k$ be a field of characteristic 0 again, let $A$ be a finite-dimensional semisimple Lie algebra over $k$, and let $V$ be a module over $A$ that is finite-dimensional as a vector space over $k$. Also let $W$ be a nonzero proper submodule of $V$. We would like to show that $V$ corresponds to a direct sum of $W$ and another submodule of $V$. This will imply Weyl’s theorem, as in Section 13.1.

Remember that the space $\mathcal{L}(V,W)$ of all linear mappings from $V$ into $W$ may be considered as a module over $A$ too, as in Section 7.5. More precisely, if $a \in A$ and $T \in \mathcal{L}(V,W)$, then $a \cdot T$ is defined as a linear mapping from $V$ into $W$ by putting

$$ (a \cdot T)(v) = a \cdot (T(v)) - T(a \cdot v) $$

for every $v \in V$. This uses the action of $A$ on $V$ in both terms on the right, and the fact that this action sends $W$ into itself, by hypothesis.

Let $Z$ be the collection of linear mappings $T$ from $V$ into $W$ for which there is an $\alpha(T) \in k$ such that

$$ T(w) = \alpha(T) w $$

for every $w \in W$. It is easy to see that $Z$ is a linear subspace of $\mathcal{L}(V,W)$, and that $\alpha$ defines a linear mapping from $Z$ onto $k$. If $a \in A$, $T \in Z$, and $w \in W$, then

$$ (a \cdot T)(w) = a \cdot (T(w)) - T(a \cdot w) = a \cdot (\alpha(T) w) - \alpha(T) (a \cdot w) = 0. $$

This implies that $a \cdot T \in Z$, with $\alpha(a \cdot T) = 0$. In particular, $Z$ is a submodule of $\mathcal{L}(V,W)$, as a module over $A$.

Let $Z_0$ be the collection of linear mappings $T$ from $V$ into $W$ such that $T \equiv 0$ on $W$. Equivalently, this means that $T \in Z$, with $\alpha(T) = 0$. Note that $Z_0$ is a submodule of $Z$, as a module over $A$. More precisely, if $a \in A$ and $T \in Z$, then
a \cdot T \in Z_0, as in the preceding paragraph. The codimension of \( Z_0 \) in \( Z \) is equal to one, because \( Z_0 \) is the kernel of \( \alpha \) on \( Z \).

The splitting theorem discussed in Section 13.1 and the previous section implies that there is a one-dimensional submodule of \( Z \), as a module over \( A \), that is complementary to \( Z_0 \). Let \( R \) be a nonzero element of this one-dimensional complementary submodule. Note that \( \alpha(R) \neq 0 \), since otherwise \( R \in Z_0 \), which would imply that \( R = 0 \). We may as well suppose that \( \alpha(R) = 1 \), by multiplying \( R \) by \( 1/\alpha(R) \). If \( a \in A \), then \( a \cdot R \in Z_0 \), as before, which implies that \( a \cdot R = 0 \), because \( a \cdot R \) is in the submodule of \( Z \) complementary to \( Z_0 \). This means that \( R \) commutes with the actions of \( A \) on \( V \) and \( W \). Thus the kernel of \( R \) is a submodule of \( V \), as a module over \( A \). The kernel of \( R \) is also complementary to \( W \) in \( V \), as desired.

### 13.7 Symmetric forms and tensors

Let \( k \) be a field, let \( A \) be a finite-dimensional vector space over \( k \), and let \( \beta(\cdot, \cdot) \) be a bilinear form on \( A \). Thus

\[
\beta_z(x) = \beta(x, z)
\]

(13.7.1)

defines a linear functional on \( A \) for each \( z \in A \), and \( z \mapsto \beta_z \) defines a linear mapping from \( A \) into the dual space \( A' \) of all linear functionals on \( A \). Let \( y, z \in A \) be given, and put

\[
T_{y,z}(x) = \beta_z(x) y = \beta(x, z) y
\]

(13.7.2)

for every \( x \in A \), which defines \( T_{y,z} \) as a linear mapping from \( A \) into itself. Observe that

\[
\beta(T_{y,z}(x), w) = \beta(x, z) \beta(y, w)
\]

(13.7.3)

and

\[
\beta(x, T_{z,y}(w)) = \beta(x, z) \beta(w, y)
\]

(13.7.4)

for every \( w, x, y, z \in A \). If \( \beta(\cdot, \cdot) \) is symmetric as a bilinear form on \( A \), then we get that

\[
\beta(T_{y,z}(x), w) = \beta(x, T_{z,y}(w))
\]

(13.7.5)

for every \( w, x, y, z \in A \).

Let us suppose from now on in this section that \( \beta(\cdot, \cdot) \) is nondegenerate on \( A \). If \( T \) is any linear mapping from \( A \) into itself, then there is a unique adjoint linear mapping \( T^* \) from \( A \) into itself such that

\[
\beta(T(x), w) = \beta(x, T^*(w))
\]

(13.7.6)

for every \( w, x \in A \), as in Section 2.14. If \( \beta(\cdot, \cdot) \) is symmetric on \( A \), then

\[
(T_{y,z})^* = T_{z,y}
\]

(13.7.7)
for every \( y, z \in A \), by (13.7.5). Remember that \( A \otimes A \) and \( A \otimes A' \) are defined as vector spaces over \( k \), as in Section 7.12. Clearly

\[
(13.7.8) \quad (y, z) \mapsto T_{y,z}
\]
defines a mapping from \( A \times A \) into the space \( \mathcal{L}(A) \) of linear mappings from \( A \) into itself that is bilinear over \( k \). This leads to a linear mapping from \( A \otimes A \) into \( \mathcal{L}(A) \), with

\[
(13.7.9) \quad y \otimes z \mapsto T_{y,z}
\]
for every \( y, z \in A \). More precisely, we have seen that there is a natural isomorphism from \( A \otimes A' \) onto \( \mathcal{L}(A) \), as vector spaces over \( k \), as in Section 13.3. Because \( \beta(\cdot, \cdot) \) is nondegenerate on \( A \), \( z \mapsto \beta_z \) is an isomorphism from \( A \) onto \( A' \), as vector spaces over \( k \). This leads to an isomorphism from \( A \otimes A \) onto \( A \otimes A' \), as vector spaces over \( k \), as before. We can compose these mappings to get an isomorphism from \( A \otimes A \) onto \( \mathcal{L}(A) \), as vector spaces over \( k \). This is the same as the mapping determined by (13.7.9), by construction.

There is a natural automorphism on \( A \otimes A \), as a vector space over \( k \), with

\[
(13.7.10) \quad y \otimes z \mapsto z \otimes y
\]
for every \( y, z \in A \), as in Section 7.15. Let us suppose from now on in this section that \( \beta(\cdot, \cdot) \) is symmetric on \( A \). Of course,

\[
(13.7.11) \quad (y, z) \mapsto T_{z,y} = (T_{y,z})^*
\]
defines a mapping from \( A \times A \) into \( \mathcal{L}(A) \), which is bilinear over \( k \). This leads to a linear mapping from \( A \otimes A \) into \( \mathcal{L}(A) \), with

\[
(13.7.12) \quad y \otimes z \mapsto T_{z,y} = (T_{y,z})^*
\]
for every \( y, z \in A \). This is the same as the composition of the mapping from \( A \otimes A \) into \( \mathcal{L}(A) \) determined by (13.7.9) with \( T \mapsto T^* \), as a linear mapping from \( \mathcal{L}(A) \) onto itself. This is also the same as the composition of the automorphism on \( A \otimes A \) determined by (13.7.10) with the linear mapping from \( A \otimes A \) into \( \mathcal{L}(A) \) determined by (13.7.9). More precisely, these linear mappings from \( A \otimes A \) into \( \mathcal{L}(A) \) are the same, because they correspond to the same bilinear mapping from \( A \times A \) into \( \mathcal{L}(A) \).

An element of \( T^2A = A \otimes A \) is said to be symmetric if it is invariant under the automorphism determined by (13.7.10), as in Section 7.15. Symmetric elements of \( A \otimes A \) correspond exactly to linear mappings from \( A \) into itself that are self-adjoint with respect to \( \beta(\cdot, \cdot) \), under the vector space isomorphism from \( A \otimes A \) onto \( \mathcal{L}(A) \) determined by (13.7.9), as in the preceding paragraph. Note that the identity mapping on \( A \) is automatically self-adjoint with respect to \( \beta(\cdot, \cdot) \). Thus the element of \( A \otimes A \) that corresponds to the identity mapping on \( A \) under the isomorphism just mentioned is symmetric in \( A \otimes A \) under these conditions. This gives another way to look at the symmetry condition mentioned at the end of Section 13.3.
13.8 Reductive Lie algebras

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. Remember that $A$ may be considered as a module over itself, using the adjoint representation. The kernel of the adjoint representation, as a Lie algebra homomorphism from $A$ into $gl(A)$, is the same as the center $Z(A)$ of $A$ as a Lie algebra. Of course, the quotient $A/Z(A)$ may be considered as a Lie algebra over $k$ too, because $Z(A)$ is an ideal in $A$. Thus $A$ may be considered as a module over $A/Z(A)$, as a Lie algebra over $k$.

An ideal in $A$ as a Lie algebra is the same as a submodule of $A$, as a module over itself with respect to the adjoint representation. Similarly, ideals in $A$ as a Lie algebra are the same as submodules of $A$ as a module over $A/Z(A)$.

Now let $k$ be a field of characteristic 0, and suppose that $A$ is a finite-dimensional Lie algebra over $k$. Suppose also that $A$ is reductive as a Lie algebra, as in Section 11.3. Thus $Z(A)$ is the same as the solvable radical of $A$. In this case, $A/Z(A)$ is semisimple as a Lie algebra over $k$, as in Section 9.4.

Of course, $Z(A)$ is a submodule of $A$, as a module over $A/Z(A)$. Because $A/Z(A)$ is semisimple as a Lie algebra over $k$, there is a submodule $B$ of $A$, as a module over $A/Z(A)$, such that $A$ corresponds to the direct sum of $Z(A)$ and $B$, as in Section 13.6. This means that

$$
Z(A) \cap B = \{0\}, \quad Z(A) + B = A,
$$

and that $B$ is an ideal in $A$, as before.

Observe that

$$
[A, A] = [B, B] \subseteq B.
$$

We also have that

$$
[A/Z(A), A/Z(A)] = A/Z(A),
$$

because $A/Z(A)$ is semisimple as a Lie algebra over $k$, as in Section 10.15. The canonical quotient mapping $q$ from $A$ onto $A/Z(A)$ maps $[A, A]$ onto the left side of (13.8.3). In this situation, the restriction of $q$ to $B$ is a one-to-one mapping from $B$ onto $A/Z(A)$. It follows that

$$
B = [A, A],
$$

and that $[A, A]$ is semisimple as a Lie algebra over $k$. This shows that $A$ corresponds to the direct sum of $Z(A)$ and $[A, A]$ as a Lie algebra over $k$, by definition of $Z(A)$. This corresponds to the second part of part (a) of Exercise 5 on p30 of [13], and to part (a) of the proposition on p102 of [13]. This also seems to correspond to a comment at the bottom of p50 in [24].

13.9 Semisimple ideals

Let $k$ be a field of characteristic 0, and let $(B, [\cdot, \cdot]_B)$ be a Lie algebra over $k$ with finite dimension as a vector space over $k$. Suppose that $A$ is an ideal in $B$
that is semisimple as a Lie algebra over $k$. Under these conditions, there is a unique ideal $B_0$ in $B$ such that $B$ corresponds to the direct sum of $A$ and $B_0$, as a Lie algebra over $k$. This is Corollary 1 on p47 of [24].

Of course, $B$ may be considered as a module over itself, using the adjoint representation. We can restrict the action of $B$ on itself to an action of $A$ on $B$, so that $B$ becomes a module over $A$, as a Lie algebra over $k$. Note that $A$ may be considered as a submodule of $B$, as a module over $A$. As in Section 13.6, there is a submodule $B_0$ of $B$, as a module over $A$, such that $B$ corresponds to the direct sum of $A$ and $B_0$, as a module over $A$. In particular, $B$ corresponds to the direct sum of $A$ and $B_0$, as a vector space over $k$, so that

$$A \cap B_0 = \{0\} \quad \text{and} \quad A + B_0 = B.$$  

(13.9.1)

Let us check that

$$[A, B_0] = \{0\},$$

(13.9.2)

where the left side is defined as in Section 9.2, as usual. Clearly

$$[A, B_0] \subseteq [A, B] \subseteq A,$$

(13.9.3)

because $A$ is an ideal in $B$. We also have that $[A, B_0] \subseteq B_0$, because $B_0$ is a submodule of $B$, as a module over $A$. Thus $[A, B_0] \subseteq A \cap B_0 = \{0\}$, as desired.

Suppose that $y \in B$ satisfies $[a, y]_B = 0$ for every $a \in A$. Because $A + B_0 = B$, there are $x \in A$ and $z \in B_0$ such that $y = x + z$. We already know that $[a, z]_B = 0$, by (13.9.2), and so we get that $[a, x]_B = 0$ for every $a \in A$. This implies that $x = 0$, because $x \in A$ and $A$ is semisimple, so that the center of $A$ is trivial. It follows that $y = z \in B_0$. This shows that $B_0$ is exactly the set of $y \in B$ such that $[a, y]_B = 0$ for every $a \in A$. Thus $B_0$ is uniquely determined by the properties of being a submodule of $B$, as a module over $A$, that is complementary to $A$.

Equivalently, $B_0$ is the centralizer of $A$ in $B$, as in Section 7.6. It is easy to see that $B_0$ is an ideal in $B$, as a Lie algebra over $k$, because $A$ is an ideal in $B$. More precisely, if $a \in A$, $w \in B$, and $y \in B_0$, then one can verify that

$$[a, [w, y]_B] = 0,$$

(13.9.4)

using the Jacobi identity and the fact that $[a, w]_B \in A$, because $A$ is an ideal in $B$. We also have that $B$ corresponds to the direct sum of $A$ and $B_0$, as a Lie algebra over $k$, because of (13.9.2).

13.10 Another approach

Let us mention another approach to the statement in the previous section, using an argument like the one in the proof of the second theorem on p23 of [13]. Let $k$ be a field, and let $(B, [\cdot, \cdot]_B)$ be a finite-dimensional Lie algebra over $k$. If $x \in B$, then

$$\text{ad}_{B, x}(z) = [x, z]_B$$

(13.10.1)
defines a linear mapping from $B$ into itself, as in Section 2.4. Put

$$b(x, y) = \text{tr}_B(\text{ad}_B x \circ \text{ad}_B y)$$

for every $x, y \in B$, which is the Killing form on $B$, as in Section 7.9. Remember that

$$b([x, w], y) = b(x, [w, y])$$

for every $w, x, y \in B$, which is to say that (13.10.2) is associative as a bilinear form on $B$, or equivalently that (13.10.2) is invariant with respect to the adjoint representation on $B$.

Let $A$ be an ideal in $B$, as a Lie algebra over $k$. If $x \in A$, then let $\text{ad}_{A,x}$ be the restriction of (13.10.1) to $z \in A$, which defines a linear mapping from $A$ into itself. Put

$$b_A(x, y) = \text{tr}_A(\text{ad}_{A,x} \circ \text{ad}_{A,y})$$

for every $x, y \in A$, which is the Killing form on $A$, as a Lie algebra over $k$. If $x, y \in A$, then

$$b_A(x, y) = b(x, y),$$

as in Section 7.10, because $A$ is an ideal in $B$.

Put

$$A^\perp = \{x \in B : b(x, y) = 0 \text{ for every } y \in A\}.$$

It is easy to see that this is an ideal in $B$, using (13.10.3) and the hypothesis that $A$ is an ideal in $B$, as in Section 10.14. Let us suppose from now on in this section that (13.10.4) is nondegenerate on $A$. Remember that this holds when $A$ is semisimple as a Lie algebra and $k$ has characteristic 0, as in Section 10.13.

It follows that

$$A \cap A^\perp = \{0\}$$

in this situation. Note that $[A, A^\perp] \subseteq A \cap A^\perp$, because $A$ and $A^\perp$ are ideals in $A$. Hence

$$[A, A^\perp] = \{0\},$$

by (13.10.7).

If $x \in B$ and $y \in A$, then put

$$\lambda_x(y) = b(x, y),$$

which defines $\lambda_x$ as a linear functional on $A$. Using the nondegeneracy of (13.10.4), we get that there is a $w \in A$ such that

$$\lambda_x(y) = b_A(w, y)$$

for every $y \in A$. This means that

$$b(x, y) = b_A(w, y) = b(w, y)$$

for every $y \in A$, so that $x - w \in A^\perp$. Thus

$$A + A^\perp = B.$$
This shows that $B$ corresponds to the direct sum of $A$ and $A^\perp$, as a Lie algebra over $k$.

One can check that $A^\perp$ is the same as the centralizer of $A$ in $B$, as in the previous section. More precisely, $A^\perp$ is contained in the centralizer of $A$, by (13.10.8). The nondegeneracy of (13.10.4) on $A$ implies that the center of $A$ is trivial, and in fact that $A$ is semisimple as a Lie algebra, as in Section 10.13. One can use this and (13.10.12) to get that $A^\perp$ is the centralizer of $A$, as before.

### 13.11 Inner derivations

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. If $x \in A$, then let $\text{ad} x$ be the mapping from $A$ into itself that sends $y \in A$ to $[x, y]_A$, as usual. Remember that the space $\text{Der}(A)$ of derivations on $A$ is a Lie algebra over $k$ with respect to the commutator bracket associated to composition of mappings, as in Section 2.5. If $x \in A$, then $\text{ad} x \in \text{Der}(A)$, and $x \mapsto \text{ad} x$ is a Lie algebra homomorphism from $A$ into $\text{Der}(A)$, as in Section 2.4. Let

\begin{equation}
\text{ad} A = \{ \text{ad} x : x \in A \}
\end{equation}

be the image of this mapping, whose elements are said to be *inner derivations* on $A$.

Let $x, y \in A$ and $\delta \in \text{Der}(A)$ be given, and observe that

\begin{equation}
([\delta, \text{ad} x])(y) = \delta((\text{ad} x)(y)) - (\text{ad} x)(\delta(y))
\end{equation}

and

\begin{equation}
\delta([x, y]_A) - [x, \delta(y)]_A = [\delta(x), y]_A = (\text{ad} \delta(x))(y).
\end{equation}

This means that

\begin{equation}
[\delta, \text{ad} x] = \text{ad} \delta(x),
\end{equation}

which implies that $\text{ad} A$ is an ideal in $\text{Der}(A)$. This corresponds to Exercise 2.1 on p9 of [13], and some remarks on p23 of [13]. This also comes up in the proof of Corollary 2 on p48 of [24].

Suppose from now on in this section that $k$ is a field, and that $A$ is a finite-dimensional semisimple Lie algebra over $k$. Remember that the kernel of the adjoint representation of $A$ is the center of $A$, which is trivial, because $A$ is semisimple. Thus the adjoint representation is a Lie algebra isomorphism from $A$ onto $\text{ad} A$, so that $\text{ad} A$ is semisimple as a Lie algebra over $k$ in particular. If $k$ has characteristic 0, then it follows that $\text{Der}(A)$ corresponds to the direct sum of $\text{ad} A$ and the centralizer of $\text{ad} A$ in $\text{Der}(A)$, as in Section 13.9. If $x \in A$ and $\delta \in \text{Der}(A)$ is in the centralizer of $\text{ad} A$ in $\text{Der}(A)$, then $\text{ad} \delta(x) = 0$, by (13.11.3). This implies that $\delta(x) = 0$, because the kernel of the adjoint representation of $A$ is trivial. This means that $\delta = 0$, so that every derivation on $A$ is an element of $\text{ad} A$. This corresponds to Corollary 2 on p48 of [24].

This also corresponds to the second theorem on p23 of [13]. The proof given there uses the argument in the previous section instead of the one in Section 13.9. This could be used when the Killing form on $A$ is nondegenerate, without asking that $k$ have characteristic 0.
Chapter 14

Nilpotence and diagonalizability

14.1 A duality argument

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [,], A)$ be a Lie algebra over $k$. If $x \in A$, then put $\text{ad}_x(y) = [x,y]_A$ for every $y \in A$, as usual. Let $\beta(\cdot, \cdot)$ be a bilinear form on $A$ that is associative or invariant under the adjoint representation on $A$, so that

$$\beta(\text{ad}_x(w), z) = -\beta(w, \text{ad}_x(z))$$ \hspace{1cm} (14.1.1)

for every $w, x, z \in A$. In particular,

$$\beta(\text{ad}_x(w), z) = 0$$ \hspace{1cm} (14.1.2)

when $[x,z]_A = 0$. This corresponds to part of part (a) of Exercise 2 on p54 of [24].

Suppose now that $k$ is a field, that $A$ has finite dimension as a vector space over $k$, and that $\beta(\cdot, \cdot)$ is also nondegenerate as a bilinear form on $A$. Let $x \in A$ be given, and note that the image

$$\{\text{ad}_x(w) : w \in A\}$$ \hspace{1cm} (14.1.3)

of $\text{ad}_x$ is a linear subspace of $A$. If $z \in A$, then (14.1.2) holds for every $w \in A$ if and only if

$$\beta(w, \text{ad}_x(z)) = 0$$ \hspace{1cm} (14.1.4)

for every $w \in A$, by (14.1.1). Because $\beta(\cdot, \cdot)$ is nondegenerate on $A$, (14.1.4) holds for every $w \in A$ if and only if $\text{ad}_x(z) = 0$. Thus (14.1.2) holds for every $w \in A$ if and only if $\text{ad}_x(z) = 0$.

Consider

$$\{y \in A : \beta(y, z) = 0 \text{ for every } z \in A \text{ such that } [x,z]_A = 0\}$$ \hspace{1cm} (14.1.5)
14.2. A CRITERION FOR NILPOTENCE

which is a linear subspace of $A$. Of course, (14.1.3) is contained in (14.1.5), because (14.1.2) holds for every $w \in A$ when $[x, z]_A = 0$, as in the preceding paragraph. In fact, (14.1.3) is equal to (14.1.5) in this situation, because $\beta(\cdot, \cdot)$ is nondegenerate on $A$. More precisely, this uses the fact that if $z \in A$ satisfies (14.1.2) for every $w \in A$, then $[x, z]_A = 0$, and hence $\beta(y, z) = 0$ for every $y$ in (14.1.5). This corresponds to the second part of part (a) of Exercise 2 on p54 of [24].

Let us now take $\beta(\cdot, \cdot)$ to be the Killing form on $A$, so that

$$\beta(u, v) = \text{tr}_A(\text{ad}_u \circ \text{ad}_v)$$

for every $u, v \in A$. Remember that this satisfies (14.1.1), as in Section 7.9.

Suppose that $x, z \in A$ commute, in the sense that $[x, z]_A = 0$. This implies that $\text{ad}_x$ and $\text{ad}_z$ commute as linear mappings from $A$ into itself, because the adjoint representation of $A$ is a representation of $A$ as a Lie algebra over $k$. If $\text{ad}_x$ is also nilpotent as a linear mapping from $A$ into itself, then it follows that $\text{ad}_x \circ \text{ad}_z$ is nilpotent as a linear mapping from $A$ into itself, as in Section 9.7.

In this case, we get that

$$\beta(x, z) = \text{tr}_A(\text{ad}_x \circ \text{ad}_z) = 0,$$

by standard arguments. If the Killing form is nondegenerate on $A$, then the previous arguments imply that there is a $w \in A$ such that

$$\text{ad}_x(w) = x.$$

This corresponds to part of part (b) of Exercise 2 on p54 of [24].

14.2 A criterion for nilpotence

Let $k$ be a field of characteristic 0, and let $A$ be an associative algebra over $k$, where multiplication of $a, b \in A$ is expressed as $ab$. Also let $\delta$ be a derivation on $A$, and suppose that $b \in A$ is an eigenvector of $\delta$ with eigenvalue $\lambda \in k$, so that

$$\delta(b) = \lambda b.$$  

In particular, this means that $b$ commutes with $\delta(b)$. If $j$ is an integer with $j \geq 2$, then

$$\delta(b^j) = j \cdot b^{j-1} \delta(b) = j \cdot \lambda b^j,$$

so that $b^j$ is an eigenvector of $\delta$ with eigenvalue $j \cdot \lambda$. If $\lambda \neq 0$, then the eigenvalues $j \cdot \lambda$ with $j \in \mathbb{Z}_+$ are all distinct, because $k$ has characteristic 0. If, for each $j \in \mathbb{Z}_+$, $b^j \neq 0$, then it follows that the $b^j$'s are linearly independent in $A$, as a vector space over $k$. If $A$ has finite dimension as a vector space over $k$, then $b^j = 0$ for some $j \geq 1$.

Remember that

$$\delta_a(x) = [a, x] = ax - xa$$

(14.2.3)
CHAPTER 14. NILPOTENCE AND DIAGONALIZABILITY

defines a derivation on $\mathcal{A}$ for every $a \in \mathcal{A}$. Suppose that $b \in \mathcal{A}$ is an eigenvector of $\delta_a$ with eigenvalue $\lambda \in \mathcal{k}$, which is to say that

$$\delta_a(b) = [a, b] = \lambda b.$$  

If $\lambda \neq 0$ and $\mathcal{A}$ has finite dimension as a vector space over $\mathcal{k}$, then $b$ is nilpotent in $\mathcal{A}$, as in the preceding paragraph.

Now let $(\mathcal{A}, [, , ]_{\mathcal{A}})$ be a Lie algebra over $\mathcal{k}$, and suppose that $u, x \in \mathcal{A}$ satisfy

$$[u, x]_{\mathcal{A}} = x.$$  

Let $V$ be a vector space over $\mathcal{k}$, and let $\rho$ be a representation of $\mathcal{A}$ as a Lie algebra on $V$. Thus

$$\rho([u, x]_{\mathcal{A}}) = [\rho u, \rho x] = \rho u \circ \rho x - \rho x \circ \rho u$$

on $V$. This means that

$$[\rho u, \rho x] = \rho x,$$

by (14.2.5). Suppose that $V$ has finite dimension, so that the algebra $\mathcal{L}(V)$ of linear mappings from $V$ into itself has finite dimension as a vector space over $\mathcal{k}$ too. Under these conditions, we get that $\rho_x$ is nilpotent as a linear mapping from $V$ into itself. This corresponds to part of part (b) of Exercise 2 on p54 of [24].

Suppose that $\mathcal{A}$ has finite dimension as a vector space over $\mathcal{k}$. Let $x$ be an ad-nilpotent element of $\mathcal{A}$, so that $\text{ad}_x$ is nilpotent on $\mathcal{A}$. If $\mathcal{A}$ is semisimple as a Lie algebra over $\mathcal{k}$, then the Killing form on $\mathcal{A}$ is nondegenerate, as in Section 10.13. In this case, there is a $u \in \mathcal{A}$ such that (14.2.5) holds, as in the previous section. This is another part of part (b) of Exercise 2 on p54 in [24].

14.3 ad-Diagonalizability

Let $\mathcal{A}$ be a Lie algebra over $\mathcal{k}$, with finite dimension as a vector space over $\mathcal{k}$. If $x \in \mathcal{A}$, then we put $\text{ad}_x(y) = [x, y]_{\mathcal{A}}$ for every $y \in \mathcal{A}$, as usual. If $\text{ad}_x$ is diagonalizable as a linear mapping from $\mathcal{A}$ into itself, then $x$ is said to be ad-diagonalizable as an element of $\mathcal{A}$. Similarly, $x$ is said to be ad-semisimple as an element of $\mathcal{A}$ when $\text{ad}_x$ is semisimple as a linear mapping from $\mathcal{A}$ into itself, as on p24 of [13], and Definition 5.5 on p52 of [24].

Suppose that $\mathcal{A}$ is an algebraically closed field of characteristic 0, and that $\mathcal{A}$ is semisimple as a Lie algebra over $\mathcal{k}$. Let $x \in \mathcal{A}$ be given, so that $\text{ad}_x$ is a linear mapping from $\mathcal{A}$ into itself. Remember that $\text{ad}_x$ can be expressed in
14.3. AD-DIAGONALIZABILITY

a unique way as a sum of two commuting linear mappings from $A$ into itself, where one of these linear mappings is diagonalizable and the other is nilpotent, as in Section 10.8. We have also seen that $\text{ad}_x$ is a derivation on $A$, as a Lie algebra over $k$. The diagonalizable and nilpotent parts of $\text{ad}_x$ are derivations on $A$ as well, as in Section 10.9. Because $k$ has characteristic 0 and $A$ is semisimple, every derivation on $A$ is an inner derivation, as in Section 13.11. This means that there are $x_1, x_2 \in A$ such that the diagonalizable and nilpotent parts of $\text{ad}_x$ are given by $\text{ad}_{x_1}$ and $\text{ad}_{x_2}$, respectively. Of course, $x_1$ and $x_2$ are uniquely determined by these properties, because of the analogous uniqueness property for the diagonalizable and nilpotent parts of $\text{ad}_x$. This corresponds to some more of the remarks on p24 of [13].

Let $V$ be a finite-dimensional vector space over $k$ again, and suppose that $A$ is a Lie subalgebra of $gl(V)$. Let $x \in A$ be given, so that $x$ is a linear mapping from $V$ into itself. As before, $x$ can be expressed in a unique way as $y_1 + y_2$, where $y_1$ and $y_2$ are commuting linear mappings from $V$ into itself, $y_1$ is diagonalizable on $V$, and $y_2$ is nilpotent on $V$. Of course, one would like to have $y_1, y_2 \in A$, as on p24 of [13]. It is easy to get this when $A = sl(V)$, as in [13]. Indeed, $y_2$ has trace 0 on $V$, because $y_2$ is nilpotent on $V$. This implies that $y_1 \in sl(V)$ in this case, because $x \in A$ by hypothesis. Some other situations will be discussed in the next section.

If $w, z \in gl(V)$, then put

\[
\text{ad}_{gl(V), w}(z) = [w, z] = w \circ z - z \circ w,
\]

as usual. Let $y \in gl(V)$ be given, so that $y$ can be expressed in a unique way as $y_1 + y_2$, where $y_1$ and $y_2$ are commuting linear mappings from $V$ into itself, $y_1$ is diagonalizable on $V$, and $y_2$ is nilpotent on $V$, as before. It follows that

\[
\text{ad}_{gl(V), y} = \text{ad}_{gl(V), y_1} + \text{ad}_{gl(V), y_2},
\]

$\text{ad}_{gl(V), y_1}$ commutes with $\text{ad}_{gl(V), y_2}$ as linear mappings on $gl(V)$, $\text{ad}_{gl(V), y_1}$ is diagonalizable on $gl(V)$, and $\text{ad}_{gl(V), y_2}$ is nilpotent on $gl(V)$, as in Section 10.9. Let $A$ be a Lie subalgebra of $gl(V)$. If $w \in A$, then $\text{ad}_{gl(V), w}$ maps $A$ into itself, and the restriction of $\text{ad}_{gl(V), w}$ to $A$ is the same as $\text{ad}_{A, w}$. If $y_1 \in A$, then it follows that $\text{ad}_{A, y_1}$ is diagonalizable as a linear mapping from $A$ into itself. Similarly, if $y_2 \in A$, then $\text{ad}_{A, y_2}$ is nilpotent as a linear mapping from $A$ into itself. Of course, $\text{ad}_{A, y_1}$ commutes with $\text{ad}_{A, y_2}$ on $A$, because $\text{ad}_{gl(V), y_1}$ commutes with $\text{ad}_{gl(V), y_2}$ on $gl(V)$. Under these conditions, we also have that $y \in A$, and

\[
\text{ad}_{A, y} = \text{ad}_{A, y_1} + \text{ad}_{A, y_2}.
\]

This corresponds to some more of the remarks on p24 of [13].
14.4 Semisimple subalgebras of $gl(V)$

Let $k$ be an algebraically closed field of characteristic 0, and let $V$ be a finite-dimensional vector space over $k$. Also let $A$ be a Lie subalgebra of $gl(V)$, and suppose that $A$ is semisimple as a Lie algebra over $k$. Let $x \in A$ be given, and remember that $x$ can be expressed in a unique way as $y_1 + y_2$, where $y_1$ and $y_2$ are commuting linear mappings from $V$ into itself, $y_1$ is diagonalizable on $V$, and $y_2$ is nilpotent on $V$. Under these conditions, $y_1, y_2 \in A$, as in the theorem on p29 of [13], and Corollary 5.4 on p52 of [24]. We shall follow the argument in [13] here, with some help from Section 13.9.

If $w \in gl(V)$, then let $\text{ad}_{gl(V),w}$ be as in (14.3.2). Thus $\text{ad}_{gl(V),w}$ maps $A$ into itself, because $x \in A$. Remember that $\text{ad}_{gl(V),y_1}$ and $\text{ad}_{gl(V),y_2}$ are the corresponding diagonalizable and nilpotent parts of $\text{ad}_{gl(V),w}$, as a linear mapping from $gl(V)$ into itself, as in Section 10.9. It follows that $\text{ad}_{gl(V),y_1}$ and $\text{ad}_{gl(V),y_2}$ map $A$ into itself as well, as in Section 10.8. Let $N = \mathcal{N}_{gl(V)}(A)$ be the normalizer of $A$ in $gl(V)$, as in Section 9.8. The previous statement says that

(14.4.1) \hspace{1cm} y_1, y_2 \in N.

If $N$ were equal to $A$, then the proof would be finished, but this does not work, because constant multiples of the identity mapping on $V$ are automatically in $N$. Thus we need some additional conditions on $y_1, y_2$.

Note that $V$ may be considered as a module over $A$, as a Lie algebra over $k$. Let $W$ be an $A$-submodule of $V$, which is to say a linear subspace of $V$ such that every element of $A$ maps $W$ into itself. Let $BW$ be the collection of $z \in gl(V)$ such that $z(W) \subseteq W$, and the restriction of $z$ to $W$ has trace equal to 0, as a linear mapping from $W$ into itself. Let us check that

(14.4.2) \hspace{1cm} A \subseteq BW.

If $z \in A$, then $z(W) \subseteq W$, by our hypothesis about $W$. Remember that $A = [A, A]$, because $A$ is semisimple and $k$ has characteristic 0, as in Section 10.15. One can use this to show that the trace of the restriction of $z$ to $W$ is 0, as desired. Note that $BW$ is a Lie subalgebra of $gl(V)$. In particular, we can take $W = V$, for which we get $BV = sl(V)$.

Let us check that

(14.4.3) \hspace{1cm} y_1, y_2 \in BW

for every $A$-submodule $W$ of $V$. Remember that $x$ maps $W$ into itself, because $x \in A$. This implies that $y_1$ and $y_2$ map $W$ into itself, as in Section 10.8. The restriction of $y_2$ to $W$ is nilpotent on $W$, because $y_2$ is nilpotent on $V$. Hence the trace of the restriction of $y_2$ to $W$ is equal to 0. The trace of the restriction of $x$ to $W$ is equal to 0 too, as in the preceding paragraph. It follows that the trace of the restriction of $y_1$ to $W$ is equal to 0 as well. This gives (14.4.3), as desired.

Let $B$ be the intersection of $N$ with all the subalgebras $BW$, over all $A$-submodules $W$ of $V$. Thus $B$ is a Lie subalgebra of $gl(V)$, and

(14.4.4) \hspace{1cm} A \subseteq B,
by (14.4.2). More precisely, $A$ is an ideal in $B$, as a Lie algebra over $k$, because $B \subseteq N$, by construction. We also have that

$$(14.4.5) \quad y_1, y_2 \in B,$$

by (14.4.1) and (14.4.3). We would like to show that $A = B$.

Note that there is a unique ideal $B_0$ in $B$ such that $B$ corresponds to the direct sum of $A$ and $B_0$, as in Section 13.9. In particular, $[A, B_0] = \{0\}$, as before.

Let $W$ be an irreducible $A$-submodule of $V$. If $z \in B_0$, then $z$ maps $W$ into itself, and $z$ commutes with every element of $A$. Because $k$ is algebraically closed, Schur’s lemma implies that $z$ is equal to a constant multiple of the identity mapping on $W$, as in Section 6.14. However, we also have that the trace of the restriction of $z$ to $W$ is equal to 0, because $z \in B$. This implies that $z \equiv 0$ on $W$, because $k$ has characteristic 0.

By Weyl’s theorem, $V$ can be expressed as the direct sum of irreducible $A$-submodules, as in Section 6.14. If $z \in B_0$, then it follows that $z \equiv 0$ on $V$, using the remarks in the preceding paragraph. This means that $B_0 = \{0\}$, so that $A = B$, as desired.

### 14.5 Tensors

Let $k$ be a field, and let $V$ be a vector space over $k$ of positive finite dimension $n$. If $p$ is a positive integer, then $T^p V$ is the $p$th tensor power of $V$, as in Section 7.15, which is to say the tensor product of $p$ copies of $V$. This is a vector space over $k$ of dimension $n^p$. More precisely, if $v_1, \ldots, v_n$ is a basis for $V$, then we can get a basis for $T^p V$ using elements of the form

$$(14.5.1) \quad v_{j_1} \otimes \cdots \otimes v_{j_p},$$

where $j_1, \ldots, j_p \in \{1, \ldots, n\}$. If $p = 0$, then one can interpret $T^0 V$ as being $k$, as a one-dimensional vector space over itself.

The dual space $V'$ of $V$ is a vector space over $k$ of dimension $n$ as well. If $q$ is a nonnegative integer, then $T^q V'$ is defined as a vector space over $k$ of dimension $n^q$, as before. Let $\lambda_1, \ldots, \lambda_q$ be $q$ linear functionals on $V$ for some $q \in \mathbb{Z}_+$, and consider

$$(14.5.2) \quad \prod_{j=1}^q \lambda_j(u_j)$$

as a $k$-valued function of $(u_1, \ldots, u_q)$ in the Cartesian product of $q$ copies of $V$. This defines a multilinear mapping from the Cartesian product of $q$ copies of $V$ into $k$, which is to say a $q$-linear form on $V$. Of course, the space of $q$-linear forms on $V$ is a vector space over $k$ with respect to pointwise addition and scalar multiplication of functions. The mapping from $(\lambda_1, \ldots, \lambda_q)$ to (14.5.2) is multilinear over $k$ as a mapping from the Cartesian product of $q$ copies of $V'$ into
the space of \( q \)-linear forms on \( V \). This leads to a linear mapping from \( T^qV' \) into the space of \( q \)-linear forms on \( V \), which is in fact a vector space isomorphism.

If \( p \) and \( q \) are nonnegative integers, then put

\[(14.5.3) \quad T^{p,q}V = T^pV \otimes T^qV',\]

which may also be denoted \( V_{p,q} \), as on p40 of [24]. This reduces to \( T^qV' \) when \( p = 0 \), and to \( T^pV \) when \( q = 0 \). Note that there is a natural isomorphism from \( T^1^1V = V \otimes V' \) onto the space \( L(V) \) of linear mappings from \( V \) into itself, as vector spaces over \( k \), as in Section 13.3. Let \( v_1, \ldots, v_n \) be a basis for \( V \) again, and let \( \mu_1, \ldots, \mu_n \) be the corresponding dual basis for \( V' \), so that \( \mu_j(v_l) \) is equal to 1 when \( j = l \) and to 0 when \( j \neq l \). We can get a basis for \( T^{p,q}V' \), as a vector space over \( k \), using elements of the form

\[(14.5.4) \quad (v_{j_1} \otimes \cdots \otimes v_{j_p}) \otimes (\mu_{l_1} \otimes \cdots \otimes \mu_{l_q}),\]

where \( j_1, \ldots, j_p, l_1, \ldots, l_q \in \{1, \ldots, n\} \).

Let \( A \) be a Lie algebra over \( k \), and let \( \rho^V \) be a representation of \( A \) on \( V \). If \( \lambda \in A \) and \( \lambda \in V' \), then put

\[(14.5.5) \quad \rho^V(a)(\lambda) = -\lambda \circ \rho^V_a,\]

which is a linear functional on \( V \). This defines \( \rho^V \) as a representation of \( A \) on \( V' \), as in Section 7.2, with \( W = k \). Equivalently, if \( a \in A \), then \( \rho^V_a \) is a linear mapping from \( V \) into itself, which leads to a dual linear mapping from \( V' \) into itself, as in Section 2.13. By construction, \( \rho^V_a \) is \(-1\) times this dual linear mapping on \( V' \).

Using \( \rho^V \) and \( \rho^V' \), we can make \( T^{p,q}V \) into a module over \( A \), as in Section 7.12. More precisely, we have \( p + q \) actions of \( A \) on \( T^{p,q}V \), using the actions of \( A \) on the individual factors on \( V \) and \( V' \) in \( T^{p,q}V \). These \( p + q \) actions of \( A \) on \( T^{p,q}V \) commute with each other, and \( T^{p,q}V \) is considered as a module over \( A \) as a Lie algebra with respect to the sum of these \( p + q \) actions, as before. If \( p = 0 \), then \( T^{p,q}V = T^qV' \) can be identified with the space of \( q \)-linear forms on \( V \), as mentioned earlier. In this case, the action of \( A \) corresponds to making the space of \( q \)-linear forms on \( V \) a module over \( A \) as a Lie algebra as in Section 7.5, using the trivial action of \( A \) on \( k \), and combining the \( q \) actions of \( A \) corresponding to the \( q \) variables of a \( q \)-linear form on \( V \).

### 14.6 Induced linear mappings

Let \( k \) be a field again, and let \( V \) be a finite-dimensional vector space over \( k \) of dimension \( n \geq 1 \). Also let \( p \) and \( q \) be nonnegative integers, at least one of which is positive. Of course, \( V \) may be considered as a module over \( gl(V) \), as a Lie algebra over \( k \). As in the previous section, \( T^{p,q}V \) may be considered as a module over \( gl(V) \) as well, by summing the actions on the various factors of \( V \) and \( V' \). More precisely, let \( A \) be a linear mapping from \( V \) into itself, and let \( A' \)
be the corresponding dual linear mapping from the dual space $V'$ into itself, as in Section 2.13. Let $w_1, \ldots, w_p \in V$ and $\lambda_1, \ldots, \lambda_q \in V'$ be given, so that

$$ (14.6.1) \quad (w_1 \otimes \cdots \otimes w_p) \otimes (\lambda_1 \otimes \cdots \otimes \lambda_q) $$

is an element of $T^{p,q}V$. If $1 \leq h \leq p$, then the action of $A$ on the $h$th factor of $V$ in $T^{p,q}V$ sends (14.6.1) to

$$ (14.6.2) \quad (w_1 \otimes \cdots \otimes w_{h-1} \otimes A(w_h) \otimes w_{h+1} \otimes \cdots \otimes w_p) \otimes (\lambda_1 \otimes \cdots \otimes \lambda_q). $$

Similarly, if $1 \leq i \leq q$, then $A'$ acts on the $i$th factor of $V'$ in $T^{p,q}V$, sending (14.6.1) to

$$ (14.6.3) \quad (w_1 \otimes \cdots \otimes w_p) \otimes (\lambda_1 \otimes \cdots \otimes \lambda_{i-1} \otimes A'(\lambda_i) \otimes \lambda_{i+1} \otimes \cdots \otimes \lambda_q). $$

If we consider $T^{p,q}V$ as a module over $gl(V)$ as in the previous section, then $A$ corresponds to a linear mapping $A_{p,q}$ from $T^{p,q}V$ into itself, as on p40 of [24]. By construction, $A_{p,q}$ sends (14.6.1) to the sum of (14.6.2) over $h = 1, \ldots, p$, minus the sum of (14.6.3) over $i = 1, \ldots, q$.

Suppose for the moment that $A$ is diagonalizable on $V$, and let $v_1, \ldots, v_n$ be a basis of $V$ consisting of eigenvectors of $A$. If $\mu_1, \ldots, \mu_n$ is the corresponding dual basis for $V'$, then one can check that $\mu_j$ is an eigenvector for $A'$ for each $j = 1, \ldots, n$. Under these conditions, the collection of elements of $T^{p,q}V$ of the form (14.5.4) is a basis for $T^{p,q}V$, as in the previous section. It is easy to see that each of these elements (14.5.4) is an eigenvector for $A_{p,q}$, so that $A_{p,q}$ is diagonalizable on $T^{p,q}V$. This corresponds to part of the proof of Lemma 6.3 on p41 of [24].

Suppose now that $A$ is nilpotent on $A$, and observe that $A'$ is nilpotent on $V'$. It is easy to see that each of the $p+q$ linear mappings on $T^{p,q}V$ corresponding to $A$ in (14.6.2) and (14.6.3) are nilpotent on $T^{p,q}V$. These $p+q$ linear mappings on $T^{p,q}V$ corresponding to $A$ also commute with each other. It follows that $A_{p,q}$ is nilpotent on $T^{p,q}$ too, because $A_{p,q}$ is defined by adding and subtracting the $p+q$ linear mappings on $T^{p,q}$, as appropriate. This corresponds to another part of the proof of Lemma 6.3 on p41 of [24].

As in the previous section, $T^{p,q}V$ is a module over $gl(V)$, as a Lie algebra over $k$, with respect to this action of $gl(V)$ on $T^{p,q}V$. Equivalently, the mapping from $A \in gl(V)$ to $A_{p,q} \in gl(T^{p,q}V)$ is a Lie algebra homomorphism. This means that $A \mapsto A_{p,q}$ is a linear mapping from $gl(V)$ into $gl(T^{p,q}V)$, and that

$$ (14.6.4) \quad [A_{p,q}, B_{p,q}] = ([A, B])_{p,q} $$

for every $A, B \in gl(V)$. In particular, if $A$ and $B$ commute as linear mappings on $V$, then $A_{p,q}$ and $B_{p,q}$ commute as linear mappings on $T^{p,q}V$. This corresponds to part of the proof of Lemma 6.3 on p41 of [24] again.

Remember that there is a natural isomorphism from $T^{1,1}V = V \otimes V'$ onto the space $\mathcal{L}(V)$ of linear mappings from $V$ into itself, as vector spaces over $k$, as in the previous section. One can check that $A_{1,1}$ corresponds to $\text{ad}_A$ on $\mathcal{L}(V)$, as in Section 10.9.
CHAPTER 14. NILPOTENCE AND DIAGONALIZABILITY

14.7 Invariant tensors

Let $k$ be a field, and let $V$ be a finite-dimensional vector space over $k$ again. Also let $p$ and $q$ be nonnegative integers, at least one of which is positive, and let $t$ be an element of $T^{p,q}V$. Consider

(14.7.1) \[ A \in gl(V) : A_{p,q}(t) = 0 \],

where $A_{p,q}$ is as in the previous section. It is easy to see that (14.7.1) is a Lie subalgebra of $gl(V)$, because $A \mapsto A_{p,q}$ is a Lie algebra homomorphism from $gl(V)$ into $gl(T^{p,q}V)$.

Suppose from now on in this section that $k$ is algebraically closed. Let $A$ be a linear mapping from $V$ into itself, and let $A_1$, $A_2$ be as in Section 10.8. Thus $A_1$ and $A_2$ are commuting linear mappings from $V$ into itself, $A = A_1 + A_2$, $A_1$ is diagonalizable on $V$, and $A_2$ is nilpotent on $V$. Let $A_{p,q}$, $(A_1)_{p,q}$, and $(A_2)_{p,q}$ be the corresponding linear mappings from $T^{p,q}V$ into itself, as in the previous section. It follows that $(A_1)_{p,q}$ and $(A_2)_{p,q}$ commute on $T^{p,q}V$,

(14.7.2) \[ A_{p,q} = (A_1)_{p,q} + (A_2)_{p,q}, \]

$(A_1)_{p,q}$ is diagonalizable on $T^{p,q}V$, and $(A_2)_{p,q}$ is nilpotent on $T^{p,q}V$, as before. This means that $(A_1)_{p,q}$ and $(A_2)_{p,q}$ are the diagonalizable and nilpotent parts $(A_{p,q})_1$ and $(A_{p,q})_2$ of $A_{p,q}$, respectively, as linear mappings from $T^{p,q}V$ into itself, by uniqueness. This corresponds to Lemma 6.3 on p41 of [24].

In particular, $(A_1)_{p,q} = (A_{p,q})_1$ and $(A_2)_{p,q} = (A_{p,q})_2$ can be expressed as polynomials in $A_{p,q}$ with no constant term, as before. If $A_{p,q}(t) = 0$, then it follows that

(14.7.3) \[ (A_1)_{p,q}(t) = (A_2)_{p,q}(t) = 0. \]

Equivalently, if $A$ is an element of (14.7.1), then $A_1$ and $A_2$ are elements of (14.7.1) as well. This argument is used in the proof of Corollary 5.4 on p52 of [24].

Let $\mathcal{A}$ be a Lie subalgebra of $gl(V)$, and suppose that $\mathcal{A}$ is semisimple as a Lie algebra over $k$. If $k$ has characteristic 0, then Theorem 5.2 on p51 of [24] says that $\mathcal{A}$ can be characterized by its tensor invariants. This means that $\mathcal{A}$ can be expressed as the intersection of subalgebras of $gl(V)$ of the form (14.7.1). Corollary 5.4 on p52 of [24] says that under these conditions, $\mathcal{A}$ contains the semisimple and nilpotent parts of its elements. Remember that another approach to the latter was discussed in Section 14.4.

Now let $\mathcal{A}$ be any finite-dimensional semisimple Lie algebra over $k$. The abstract Jordan decomposition for elements of $\mathcal{A}$ is given in Theorem 5.6 on p52 of [24]. This is obtained from Corollary 5.4 in [24], applied to the image of $\mathcal{A}$ under the adjoint representation.

14.8 Diagonalizability and quotients

Let $k$ be a field, let $V$ be a vector space over $k$, and let $W$ be a linear subspace of $V$. Consider the collection $\mathcal{L}_W(V)$ of all linear mappings $T$ from $V$ into itself
such that
\begin{equation}
T(W) \subseteq W.
\end{equation}
This is a subalgebra of the algebra \(L(V)\) of all linear mappings from \(V\) into itself. Let \(q\) be the canonical quotient mapping from \(V\) onto the quotient vector space \(V/W\). If \(T \in L_W(V)\), then there is a unique linear mapping \(T^W/V\) from \(V/W\) into itself such that
\begin{equation}
T^W/V(q(v)) = q(T(v))
\end{equation}
for every \(v \in V\), as usual. It is easy to see that \(T \mapsto T^W/V\) is an algebra homomorphism from \(L_W(V)\) into \(L(V/W)\), with respect to composition of mappings. More precisely, this homomorphism maps \(L_W(V)\) onto \(L(V/W)\). We shall normally be concerned with finite-dimensional vector spaces here, in which case the previous statement is more elementary.

In particular, if \(T \in L_W(V)\) is nilpotent as a linear mapping on \(V\), then \(T^W/V\) is nilpotent as a linear mapping on \(V/W\). If \(R, T \in L_W(V)\) commute as linear mappings on \(V\), then \(R^W/V, T^W/V\) commute as linear mappings on \(V/W\).

Let \(T\) be an element of \(L_W(V)\), and suppose that \(v \in V\) is an eigenvector of \(T\), with eigenvalue \(\lambda\). Observe that
\begin{equation}
T^W/V(q(v)) = q(T(v)) = q(\lambda v) = \lambda q(v),
\end{equation}
so that \(q(v)\) is an eigenvector of \(T^W/V\) with eigenvalue \(\lambda\) as well. If \(V\) is spanned by the eigenvectors of \(T\), then it follows that \(V/W\) is spanned by the images under \(q\) of the eigenvectors of \(T\), which are eigenvectors of \(T^W/V\).

Let \((A, [\cdot, \cdot]_A)\) and \((B, [\cdot, \cdot]_B)\) be Lie algebras over \(k\). Put
\begin{equation}
ad_{A,x}(y) = [x, y]_A, \quad ad_{B,w}(z) = [w, z]_B
\end{equation}
for every \(x, y \in A\) and \(w, z \in B\), as usual. Suppose that \(\phi\) is a Lie algebra homomorphism from \(A\) into \(B\). This implies that
\begin{equation}
\phi(ad_{A,x}(y)) = \phi([x, y]_A) = [\phi(x), \phi(y)]_B = ad_{B, \phi(x)}(\phi(y))
\end{equation}
for every \(x, y \in A\). Suppose for the moment that \(\phi\) maps \(A\) onto \(B\). If \(x \in A\) and \(ad_{A,x}\) is nilpotent as a linear mapping from \(A\) into itself, then it follows that \(ad_{B, \phi(x)}\) is nilpotent as a linear mapping from \(B\) into itself. Equivalently, if \(x \in A\) is ad-nilpotent in \(A\), then \(\phi(x)\) is ad-nilpotent in \(B\).

If \(x, y \in A\) and \(y\) is an eigenvector of \(ad_{A,x}\), then \(\phi(y)\) is an eigenvector of \(ad_{B, \phi(x)}\), by (14.8.5). If \(A\) is spanned by the eigenvectors of \(ad_{A,x}\) and \(\phi\) maps \(A\) onto \(B\), then it follows that \(B\) is spanned by the eigenvectors of \(ad_{B, \phi(x)}\). These remarks correspond to part of the proof of the corollary on p30 of [13]. This also corresponds to Theorem 5.7 on p52 of [24], in the simpler case of surjective Lie algebra homomorphisms.
14.9 Diagonalizability and representations

Let $k$ be an algebraically closed field of characteristic 0, and let $V$ be a finite-dimensional vector space over $k$. Also let $B$ be a Lie subalgebra of $gl(V)$, and suppose that $B$ is semisimple as a Lie algebra over $k$. Let $x \in B$ be given, and remember that there are unique commuting linear mappings $y_1$ and $y_2$ on $V$ such that $x = y_1 + y_2$, $y_1$ is diagonalizable on $V$, and $y_2$ is nilpotent on $V$, as in Section 10.8. In fact, $y_1, y_2 \in B$, because $B$ is semisimple as a Lie algebra, as in Sections 14.4 and 14.7. Remember that $y_1$ is ad-diagonalizable on $gl(V)$, and $y_2$ is ad-nilpotent on $gl(V)$, as in Section 10.9. Using this, one can check that $y_1$ is ad-diagonalizable on $B$, and that $y_2$ is ad-nilpotent on $B$, because $B$ is a Lie subalgebra of $gl(V)$. More precisely, if $y \in B$, then $ad_B y$ is the same as the restriction of $ad_{gl(V)} y$ to $B$, because $B$ is a Lie subalgebra of $gl(V)$. This means that $y_1$ and $y_2$ are the same as the ad-diagonalizable and ad-nilpotent parts of the abstract Jordan decomposition of $x$ in $B$, by the uniqueness of the abstract Jordan decomposition, as in Section 14.3. This is the second part of the theorem on p29 of [13].

Let $(A, [\cdot, \cdot]_A)$ be a finite-dimensional semisimple Lie algebra over $k$, and let $\phi$ be a Lie algebra homomorphism from $A$ into $gl(V)$. Put $B = \phi(A)$, which is a Lie subalgebra of $gl(V)$. Note that $B$ is semisimple as a Lie algebra over $k$ too, as in Section 10.14. Let $w \in A$ be given, and let $w = w_1 + w_2$ be the abstract Jordan decomposition of $w$ in $A$, as in Section 14.3. Thus $w_1$ is ad-diagonalizable on $A$, $w_2$ is ad-nilpotent on $A$, and $[w_1, w_2]_A = 0$, as before. It follows that $\phi(w) = \phi(w_1) + \phi(w_2)$, where $\phi(w_1)$ is ad-diagonalizable on $B$ and $\phi(w_2)$ is ad-nilpotent on $B$, as in the previous section. Note that $\phi(w_1)$ and $\phi(w_2)$ commute as linear mappings on $V$, because $[w_1, w_2]_A = 0$ and $\phi$ is a Lie algebra homomorphism. This means that $\phi(w_1)$ and $\phi(w_2)$ satisfy the requirements of the abstract Jordan decomposition of $\phi(w)$ in $B$. Put $x = \phi(w)$, and let $y_1, y_2 \in B$ be as in the preceding paragraph. It follows that

$$\phi(w_1) = y_1, \quad \phi(w_2) = y_2,$$

by uniqueness of the abstract Jordan decomposition of $x$ in $B$. This corresponds to the corollary on p30 of [13], and to Theorem 7 on p27 of [23].

Let $(A_2, [\cdot, \cdot]_{A_2})$ be a finite-dimensional semisimple Lie algebra over $k$, and let $A_1$ be a Lie subalgebra of $A_2$ that is semisimple as a Lie algebra over $k$ as well. If $x, y \in A_2$, then put $ad_{A_2,x} (y) = [x, y]_{A_2}$, as usual. If $x \in A_1$, then $ad_{A_1,x}$ is the same as the restriction of $ad_{A_2,x}$ to $A_1$, because $A_1$ is a Lie subalgebra of $A_2$. Of course, $A_2$ is a finite-dimensional vector space over $k$ in particular, so that $gl(A_2)$ is a Lie algebra over $k$ with respect to the commutator bracket associated to composition of linear mappings. If $x \in A_1$, then put $\phi(x) = ad_{A_2,x}$, which defines a Lie algebra homomorphism from $A_1$ into $gl(A_2)$. Let $w \in A_1$ be given, and let $w = w' + w''$ be the abstract Jordan decomposition of $w$ in $A_1$. Thus

$$ad_{A_2,w} = ad_{A_2,w'} + ad_{A_2,w''},$$

corresponds to the ordinary Jordan decomposition of $ad_{A_2,w}$, as a linear mapping from $A_2$ into itself, as before. This means that $w = w' + w''$ also corresponds
Let $A_1, A_2$ be finite-dimensional semisimple Lie algebras over $k$, and let $\psi$ be a Lie algebra homomorphism from $A_1$ into $A_2$. Theorem 5.7 on p52 of [24] basically says that $\psi$ maps abstract Jordan decompositions in $A_1$ to the corresponding abstract Jordan decompositions in $A_2$. Of course, this includes the case of subalgebras, as in the preceding paragraph. If $\psi$ maps $A_1$ onto $A_2$, then this can be verified directly, as in the previous section. Otherwise, one can use this to reduce to the case of subalgebras.

**14.10 Exponentiating nilpotents**

Let $T_1$ and $T_2$ be commuting indeterminates, and remember that $\mathbb{Z}[T_1, T_2]$ is the space of formal polynomials in $T_1$ and $T_2$ with coefficients in $\mathbb{Z}$, as in Section 5.8. This is a commutative associative algebra over $\mathbb{Z}$ with respect to formal multiplication of polynomials, as before. If $n$ is a nonnegative integer, then

\[
(T_1 + T_2)^n = \sum_{j=0}^{n} \binom{n}{j} T_1^j T_2^{n-j},
\]

where $\binom{n}{j} = \frac{n!}{j!(n-j)!}$ is the usual binomial coefficient, as in the binomial theorem. Thus

\[
(T_1 + T_2)^{n+1} = \sum_{j=0}^{n+1} \binom{n+1}{j} T_1^j T_2^{n+1-j}
\]

is the same as

\[
(T_1 + T_2)^n (T_1 + T_2) = \sum_{j=0}^{n} \binom{n}{j} T_1^{j+1} T_2^{n-j} + \sum_{j=0}^{n} \binom{n}{j} T_1^j T_2^{n-j+1}
\]

\[
= \sum_{j=1}^{n+1} \binom{n}{j-1} T_1^j T_2^{n+1-j} + \sum_{j=0}^{n} \binom{n}{j} T_1^j T_2^{n+1-j}.
\]

This implies the well-known identity

\[
\binom{n+1}{j} = \binom{n}{j-1} + \binom{n}{j}
\]

for $j = 1, \ldots, n$.

Let $k$ be a field of characteristic 0. Alternatively, one can consider commutative rings $k$ with multiplicative identity elements which are algebras over $\mathbb{Q}$. In this case, a module or algebra over $k$ could also be considered as a vector space over $\mathbb{Q}$. Let $\mathcal{A}$ be an associative algebra over $k$ with a multiplicative identity element $e$. Suppose that $x \in \mathcal{A}$ is nilpotent, so that $x^j = 0$ for some positive
integer \( l \). Under these conditions, the exponential of \( x \) is defined as an element of \( A \) by

\[
\exp x = \sum_{j=0}^{\infty} \frac{1}{j!} x^j,
\]

as usual. More precisely, \( x^j = 0 \) for all sufficiently large \( j \), so that the infinite series reduces to a finite sum, and thus defines an element of \( A \). Of course, \( x^j \) is interpreted as being equal to \( e \) when \( j = 0 \).

Let \( y \) be another nilpotent element of \( A \), and suppose that \( x \) commutes with \( y \). This implies that \( x + y \) is nilpotent in \( A \), so that \( \exp(x + y) \) can be defined as before. Observe that

\[
\exp(x + y) = \sum_{n=0}^{\infty} \frac{1}{n!} (x + y)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{j=0}^{n} \binom{n}{j} x^j y^{n-j} = \sum_{n=0}^{\infty} \sum_{j=0}^{n} \frac{1}{j!} x^j \left(\frac{1}{(n-j)!}\right) y^{n-j} = (\exp x) (\exp y).
\]

More precisely, the double sum is the same as the Cauchy product of the series defining \( \exp x \) and \( \exp y \), as in Section 4.2. The nilpotency conditions on \( x \) and \( y \) ensure that all of these sums reduce to finite sums. In particular, we can take \( y = -x \), to get that

\[
(\exp x) \exp(-x) = \exp(-x) (\exp x) = \exp 0 = e.
\]

This means that \( \exp x \) is invertible in \( A \), with

\[
(\exp x)^{-1} = \exp(-x).
\]

Let \( V \) be a module over \( k \), and let \( T \) be a mapping from \( V \) into itself that is linear over \( k \). Suppose that \( T \) is nilpotent on \( V \), so that the \( l \)th power \( T^l \) of \( T \) with respect to composition is equal to 0 on \( V \) for some positive integer \( l \). Thus the exponential \( \exp T \) of \( T \) can be defined as a linear mapping from \( V \) into itself, using the algebra of linear mappings from \( V \) into itself with respect to composition of mappings in the previous remarks. Let \( W \) be a submodule of \( V \), and suppose that \( T \) maps \( W \) into itself. It follows that \( \exp T(W) \subseteq W \) for every nonnegative integer \( j \), so that

\[
(\exp(T))(W) \subseteq W.
\]

More precisely, we have that

\[
(\exp T)(W) = W
\]

in this situation. This can be obtained from (14.10.9) and the analogous statement for \( -T \). Alternatively, the restriction of \( \exp T \) to \( W \) is the same as the exponential of the restriction of \( T \) to \( W \), which is invertible on \( W \).
14.11. NILPOTENT DERIVATIONS

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an algebra over $k$ in the strict sense, where multiplication of $a, b \in A$ is expressed as $ab$. If $\delta$ is a derivation on $A$, then it is well known that

$$
\delta^n(ab) = \sum_{j=0}^{n} \binom{n}{j} \cdot \delta^j(a) \delta^{n-j}(b)
$$

for every nonnegative integer $n$ and $a, b \in A$. More precisely, if this holds for some $n \geq 0$, then

$$
\delta^{n+1}(ab) = \delta(\delta^n(ab)) = \sum_{j=0}^{n} \binom{n}{j} \delta^{j+1}(a) \delta^{n-j}(b) + \delta^j(a) \delta^{n-j+1}(b)
$$

This implies the analogue of (14.10.11) for $n + 1$, by (14.10.4). It follows that (14.10.11) holds for all $n \geq 0$, by induction.

14.11 Nilpotent derivations

Let $k$ be a commutative ring with a multiplicative identity element that is an algebra over $\mathbb{Q}$, or simply a field of characteristic 0, as in the previous section. Also let $A$ be an algebra over $k$ in the strict sense again, and let $\delta$ be a derivation on $A$ that is nilpotent as a mapping from $A$ into itself, so that $\delta^l = 0$ for some positive integer $l$. Thus $\exp \delta$ can be defined as a mapping from $A$ into itself that is linear over $k$, as in the previous section. If $a, b \in A$, then

$$
\exp \delta(a)(b) = \sum_{n=0}^{\infty} \frac{1}{n!} \delta^n(ab) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{j=0}^{n} \binom{n}{j} \delta^j(a) \delta^{n-j}(b)
$$

This implies that $\exp \delta$ is an algebra automorphism on $A$, as on p9 of \cite{13}. Let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$, and put $\text{ad}_x(y) = [x, y]_A$ for every $x, y \in A$, as usual. Remember that $\text{ad}_x$ is a derivation on $A$, as in Section 2.5. If $x \in A$ is ad-nilpotent, then $\exp \text{ad}_x$ defines a Lie algebra automorphism on $A$, as in the preceding paragraph. Let $\text{Int} A$ be the subgroup of the group of all Lie algebra automorphisms of $A$ generated by these automorphisms. The elements
of $\text{Int} \ A$ are called *inner automorphisms* of $A$, as on p9 of [13]. If $\phi$ is any Lie algebra automorphism on $A$, then it is easy to see that
\begin{equation}
\phi \circ \text{ad}_x \circ \phi^{-1} = \text{ad}_{\phi(x)}
\end{equation}
for every $x \in A$. In particular, if $x \in A$ is ad-nilpotent, then $\phi(x)$ is ad-nilpotent as well. In this case, we get that
\begin{equation}
\phi \circ (\exp \text{ad}_x) \circ \phi^{-1} = \exp \text{ad}_{\phi(x)},
\end{equation}
by (14.11.2). This implies that $\text{Int} \ A$ is a normal subgroup of the group of all Lie algebra automorphisms on $A$, as on p9 of [13].

Let $B$ be a Lie subalgebra of $A$, and let $x$ be an element of the normalizer of $B$ in $A$, so that $\text{ad}_x$ maps $B$ into itself. If $x$ is ad-nilpotent on $A$, then $\exp \text{ad}_x$ can be defined as a linear mapping from $A$ into itself as before. In fact, we have that
\begin{equation}
(\exp \text{ad}_x)(B) = B,
\end{equation}
as in (14.10.10). In particular, this holds when $x \in B$, in which case the restriction of $\text{ad}_x$ to $B$ is the same as $\text{ad}_{x,B}$.

Now let $A$ be an associative algebra over $k$ with a multiplicative identity element $e$, where multiplication of $a, b \in A$ is expressed as $a b$ again. If $a \in A$, then let $L_a$ and $R_a$ be the usual operators of left and right multiplication by $a$ on $A$, so that $L_a(x) = a x$ and $R_a(x) = x a$ for every $a \in A$. Note that
\begin{equation}
(L_a)^j = L_{a^j}, \quad (R_a)^j = R_{a^j}
\end{equation}
for every nonnegative integer $j$. Suppose that $a$ is nilpotent in $A$, which implies that $L_a$ and $R_a$ are nilpotent as linear mappings from $A$ into itself. It is easy to see that
\begin{equation}
\exp L_a = L_{\exp a}, \quad \exp R_a = R_{\exp a}
\end{equation}
under these conditions.

Remember that $A$ may be considered as a Lie algebra over $k$ with respect to the corresponding commutator bracket $[a, b] = a b - b a$. If $a \in A$, then
\begin{equation}
\text{ad}_a = L_a - R_a = L_a + R_{-a},
\end{equation}
as a linear mapping from $A$ into itself. Suppose that $a$ is nilpotent in $A$ again, and remember that $L_a$ and $R_{-a}$ commute as linear mappings from $A$ into itself. This implies that $\text{ad}_a$ is nilpotent on $A$, and that
\begin{equation}
\exp \text{ad}_a = \exp(L_a + R_{-a}) = (\exp L_a) \circ (\exp R_{-a}) = L_{\exp a} \circ R_{\exp(-a)} = L_{\exp a} \circ R_{\exp(-a)} \circ 1.
\end{equation}
Equivalently,
\begin{equation}
(\exp \text{ad}_a)(x) = (\exp a) x (\exp a)^{-1}
\end{equation}
for every $x \in A$, as on p9 of [13].
14.12 Exponentials and homomorphisms

Let $k$ be a commutative ring with a multiplicative identity element that is an algebra over $\mathbb{Q}$, or simply a field of characteristic $0$, as before. Also let $\mathcal{A}$ and $\mathcal{B}$ be associative algebras over $k$ with multiplicative identity elements $e_{\mathcal{A}}$ and $e_{\mathcal{B}}$, respectively. Suppose that $\phi$ is an algebra homomorphism from $\mathcal{A}$ into $\mathcal{B}$ such that $\phi(e_{\mathcal{A}}) = e_{\mathcal{B}}$. Let $x$ be a nilpotent element of $\mathcal{A}$, which implies that $\phi(x)$ is nilpotent in $\mathcal{B}$. Under these conditions, it is easy to see that

$$\phi(\exp_{\mathcal{A}} x) = \exp_{\mathcal{B}} \phi(x),$$

where $\exp_{\mathcal{A}} x$ and $\exp_{\mathcal{B}} \phi(x)$ are the exponentials of $x$ and $\phi(x)$ in $\mathcal{A}$ and $\mathcal{B}$, respectively.

Now let $(\mathcal{A}, [\cdot, \cdot]_{\mathcal{A}})$ and $(\mathcal{B}, [\cdot, \cdot]_{\mathcal{B}})$ be Lie algebras over $k$, and let $\psi$ be a Lie algebra homomorphism from $\mathcal{A}$ into $\mathcal{B}$. As usual, we put $\text{ad}_{\mathcal{A},z}(y) = [x, y]_{\mathcal{A}}$ and $\text{ad}_{\mathcal{B},w}(z) = [w, z]_{\mathcal{B}}$ for every $x, y \in \mathcal{A}$ and $w, z \in \mathcal{B}$. If $x, y \in \mathcal{A}$, then

$$\psi(\text{ad}_{\mathcal{A},x}(y)) = \text{ad}_{\mathcal{B},\phi(x)}(\psi(y)),$$

as in Section 14.8. It follows that

$$\psi((\text{ad}_{\mathcal{A},x})^j(y)) = (\text{ad}_{\mathcal{B},\phi(x)})^j(\psi(y))$$

for every positive integer $j$. Suppose that $x$ is ad-nilpotent in $\mathcal{A}$, and that $\psi(x)$ is ad-nilpotent in $\mathcal{B}$. Thus $\exp \text{ad}_{\mathcal{A},x}$ and $\exp \text{ad}_{\mathcal{B},\psi(x)}$ can be defined as linear mappings from $\mathcal{A}$ and $\mathcal{B}$ into themselves, as before. Using (14.12.3), we get that

$$\psi((\exp \text{ad}_{\mathcal{A},x})(y)) = (\exp \text{ad}_{\mathcal{B},\psi(x)})(\psi(y))$$

for every $y \in \mathcal{A}$.

If $x$ is ad-nilpotent in $\mathcal{A}$, and $\psi$ maps $\mathcal{A}$ onto $\mathcal{B}$, then $\psi(x)$ is ad-nilpotent in $\mathcal{B}$, by (14.12.3), as in Section 14.8. Similarly, if $\psi(x)$ is ad-nilpotent in $\mathcal{B}$, and $\psi$ is injective, then $x$ is ad-nilpotent in $\mathcal{A}$.

Let $V$ be a module over $k$, and let $\rho$ be a representation of $\mathcal{A}$ on $V$. If $R$ and $T$ are mappings from $V$ into itself that are linear over $k$, then put $\text{ad}_{R, T} = [R, T] = R \circ T - T \circ R$, as usual. Suppose that $x \in \mathcal{A}$ has the property that $\rho_x$ is nilpotent as a linear mapping from $V$ into itself. Thus $\exp \rho_x$ is defined as an invertible linear mapping from $V$ onto itself, as before. We also have that $\text{ad}_{\rho_x}$ is nilpotent as a linear mapping on the space of linear mappings from $V$ into itself, so that $\exp \text{ad}_{\rho_x}$ is defined as an invertible linear mapping on the space of linear mappings from $V$ into itself. In fact,

$$(\exp \text{ad}_{\rho_x})(T) = (\exp \rho_x) \circ T \circ (\exp \rho_x)^{-1}$$

for every linear mapping $T$ from $V$ into itself, as in the previous section. In particular,

$$(\exp \text{ad}_{\rho_x})(\rho_y) = (\exp \rho_x) \circ \rho_y \circ (\exp \rho_x)^{-1}$$

for every $y \in \mathcal{A}$. 
Suppose that $x$ is ad-nilpotent in $A$, in addition to $x$ being nilpotent on $V$. Thus $\exp \text{ad}_{A;x}$ is defined as a linear mapping from $A$ into itself, and

$$
(14.12.7) \quad \rho(\exp \text{ad}_{A;x}(y)) = (\exp \text{ad}_{\rho_x})(\rho y)
$$

for every $y \in A$. This follows from (14.12.4), with $B$ taken to be the space of linear mappings from $V$ into itself, considered as a Lie algebra over $k$ with respect to the commutator bracket associated to composition of linear mappings, and with $\psi$ taken to be the Lie algebra homomorphism from $A$ into $B$ corresponding to $\rho$. This implies that

$$
(14.12.8) \quad \rho(\exp \text{ad}_{A;x}(y)) = (\exp \rho_x) \circ \rho y \circ (\exp \rho_x)^{-1}
$$

for every $y \in A$.

### 14.13 Representations and structure constants

Let $k$ be a commutative ring with a multiplicative identity element, and let $m$ and $n$ be positive integers. The spaces $k^m$ and $k^n$ of $m$ and $n$-tuples of elements of $k$ are (free) modules over $k$ with respect to coordinatewise addition and scalar multiplication, as usual. Let $\rho$ be a bilinear action of $k^m$ on $k^n$, which corresponds to a mapping from $k^m \times k^n$ into $k^n$ that is bilinear over $k$. If $x = (x_1, \ldots, x_m) \in k^m$ and $v = (v_1, \ldots, v_n) \in k^n$, then $\rho_x(v) \in k^n$ can be expressed as

$$
(14.13.1) \quad (\rho_x(v))_r = \sum_{j=1}^{m} \sum_{l=1}^{n} a_{j,l}^{r} x_j v_l,
$$

where the left side is the $r$th coordinate of $\rho_x(v)$ for each $r = 1, \ldots, n$. The coefficients $a_{j,l}^{r}$ are elements of $k$ for each $j = 1, \ldots, m$ and $l, r = 1, \ldots, n$, and do not depend on $x$ or $v$. It is easy to see that these coefficients are uniquely determined by $\rho$. Conversely, any coefficients $a_{j,l}^{r} \in k$ determine a bilinear action of $k^m$ on $k^n$ in this way.

Let $K$ be a commutative associative algebra over $k$, where multiplication of $y, z \in K$ is expressed as $yz$. The spaces $K^m$ and $K^n$ of $m$ and $n$-tuples of elements of $K$ may be considered as modules over $k$ with respect to coordinatewise addition and scalar multiplication. If $x \in K^m$ and $v \in K^n$, then let $\rho^K_x(v)$ be the element of $K^n$ whose $r$th coordinate is equal to

$$
(14.13.2) \quad (\rho^K_x(v))_r = \sum_{j=1}^{m} \sum_{l=1}^{n} a_{j,l}^{r} x_j v_l
$$

for each $r = 1, \ldots, n$. This uses both multiplication on $K$ and scalar multiplication of elements of $K$ by elements of $k$ to define the terms of the sum on the right. This defines $\rho^K$ as a bilinear action of $K^m$ on $K^n$.

Suppose that $k^m$ is a Lie algebra over $k$, with respect to a Lie bracket $[\cdot, \cdot]_{k^m}$. Using the structure constants for $[\cdot, \cdot]_{k^m}$, we get a Lie bracket $[\cdot, \cdot]_{K^m}$ on $K^m$, as in Section 9.14.
Suppose also that \( \rho \) is a representation of \( k^m \), as a Lie algebra over \( k \), on \( k^n \). This means that

\[
(14.13.3) \quad \rho_{[w, x]_k} (v) = \rho_w(\rho_x(v)) - \rho_x(\rho_w(v))
\]

for every \( w, x \in k^m \) and \( v \in k^n \). This can be characterized by suitable conditions on the coefficients \( a'_{i,j} \) associated to \( \rho \) and the structure constants for \( [\cdot, \cdot]_k \). It follows that

\[
(14.13.4) \quad \rho^K_{[w, x]_k} (v) = \rho^K_w(\rho^K_x(v)) - \rho^K_x(\rho^K_w(v))
\]

for every \( w, x \in k^m \) and \( v \in K^n \). Thus \( \rho^K \) is a representation of \( K^m \), as a Lie algebra over \( k \), on \( K^n \) under these conditions.

If \( T \) is a homomorphism from \( k^n \) into itself, as a module over \( k \), then \( T \) corresponds to an \( n \times n \) matrix with entries in \( k \) in the usual way. Using the same matrix, we get a homomorphism \( T_K \) from \( K^n \) into itself, as a module over \( K \). This defines a homomorphism from \( \text{Hom}_k(k^n, k^n) \) into \( \text{Hom}_k(K^n, K^n) \), as associative algebras over \( k \) with respect to composition of mappings.

Suppose that \( K \) has a multiplicative identity element \( e \), so that \( K^m \) and \( K^n \) may be considered as free modules over \( K \) with respect to coordinatewise addition and scalar multiplication. In this case, \( K^m \) may be considered as a Lie algebra over \( K \) with respect to \( [\cdot, \cdot]_K \), as in Section 9.14. Similarly, \( \rho^K \) may be considered as a representation of \( K^m \), as a Lie algebra over \( K \), on \( K^n \), as a module over \( K \). If \( T \) is a homomorphism from \( k^n \) into itself as a module over \( k \), then \( T_K \) is a homomorphism from \( K^n \) into itself as a module over \( K \).

As before, \( t \mapsto te \) defines a ring homomorphism from \( k \) into \( K \). This leads to homomorphisms from \( k^m \) and \( k^n \) into \( K^m \) and \( K^n \), as modules over \( k \). More precisely, this gives a homomorphism from \( k^m \) into \( K^m \) as Lie algebras over \( K \). Similarly, if \( x \in k^m \) and \( v \in k^n \), then the image of \( \rho_x(v) \) in \( K^n \) is the same as taking the images of \( x \) and \( v \) in \( k^m \) and \( k^n \), respectively, and then using \( \rho^K \).

If \( T \) is a homomorphism from \( k^n \) into itself, as a module over \( k \), and \( v \in k^n \), then the image of \( T(v) \) in \( K^n \) is the same as first taking the image of \( v \) in \( k^n \), and then taking the image of that under \( T_K \).

Suppose that \( t \mapsto te \) is injective as a mapping from \( k \) into \( K \), so that the corresponding mappings from \( k^m \) and \( k^n \) into \( K^m \) and \( K^n \), respectively, are injective as well. If \( T \) is a homomorphism from \( k^n \) into itself, as a module over \( k \), and \( T_K = 0 \) on \( K^n \), then \( T = 0 \) on \( k^n \). Similarly, if \( T_K \) is nilpotent on \( K^n \), then \( T \) is nilpotent on \( k^n \).

Let \( x \in k^m \) be given, and let \( x_K \) be its image in \( K^m \). If \( T = \rho_x \), then \( T_K = \rho^K_x \). In particular, if \( \rho^K_x \) is nilpotent as a mapping from \( K^m \) into itself, then \( \rho_x \) is nilpotent on \( k^m \).

### 14.14 Lie’s theorem and nilpotence

Let \( k \) be a field of characteristic 0, and let \((A, [\cdot, \cdot]_A)\) be a solvable Lie algebra over \( k \) with positive finite dimension \( m \) as a vector space over \( k \). Also let \( V \) be a vector space over \( k \) of positive finite dimension \( n \), and let \( \rho \) be a representation of \( A \) on \( V \). Remember that \([A, A] \subseteq A\) as defined in Section 9.2.
Suppose for the moment that $k$ is algebraically closed. Under these conditions, there is a flag $\mathcal{F} = \{V_j\}_{j=0}^{n}$ in $V$ such that

\begin{equation}
(14.14.1) \quad \rho_a(V_j) \subseteq V_j
\end{equation}

for every $a \in A$ and $j = 0, 1, \ldots, n$, as in Section 9.13. If $a \in [A, A]$, then it follows that

\begin{equation}
(14.14.2) \quad \rho_a(V_j) \subseteq V_{j-1}
\end{equation}

for every $j = 1, \ldots, n$, as in Section 9.11. In particular, this implies that

\begin{equation}
(14.14.3) \quad (\rho_a)^n = 0
\end{equation}

on $V$ for every $a \in [A, A]$.

Otherwise, let $K$ be an algebraically closed field that contains $k$ as a subfield. We may as well suppose that $A = k^m$ as a vector space over $k$, and that $V = k^n$. As in Section 9.14, $[,]_A$ can be extended to a Lie bracket $[,]_K$ on $A_K = K^m$, so that $A_K$ becomes a solvable Lie algebra over $K$. Similarly, we can use $\rho$ to get a representation $\rho^K$ of $A_K$, as a Lie algebra over $K$, on $K^n$, as a vector space over $K$, as in the previous section.

Of course, $[,]_K$ can be defined as a subset of $A_K$ as in Section 9.2 too. If $a \in [A_K, A_K]$, then it follows that

\begin{equation}
(14.14.4) \quad (\rho^K_a)^n = 0
\end{equation}

on $K^n$, as before. In particular, this holds when $a \in [A, A]$. This implies that (14.14.3) holds on $V = k^n$ when $a \in [A, A]$.
Chapter 15

Complexifications and $sl_2(k)$ modules

15.1 Basic properties of $sl_2(k)$ modules

Let $k$ be a commutative ring with a multiplicative identity element. Remember that $sl_2(k)$ is freely generated, as a module over $k$, by

$$
\begin{align*}
  x &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\
  y &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \\
  h &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},
\end{align*}
$$

as in Section 10.2. Thus every element of $sl_2(k)$ can be expressed in a unique way as a linear combination of $x$, $y$, and $h$, with coefficients in $k$. We also have that

$$
\begin{align*}
  [h, x] &= 2 \cdot x, \\
  [h, y] &= 2 \cdot y, \\
  [x, y] &= h,
\end{align*}
$$

as before, using the commutator bracket associated to matrix multiplication.

Let $V$ be a module over $k$, and let $\rho = \rho^V$ be a representation of $sl_2(k)$ on $V$, as a Lie algebra over $k$. Thus $\rho_x$, $\rho_y$, and $\rho_h$ are homomorphisms from $V$ into itself, as a module over $k$, such that

$$
\begin{align*}
  [\rho_h, \rho_x] &= 2 \cdot \rho_x, \\
  [\rho_h, \rho_y] &= 2 \cdot \rho_y, \\
  [\rho_x, \rho_y] &= \rho_{[x, y]} = \rho_h,
\end{align*}
$$

Equivalently, we may consider $V$ as a module over $sl_2(k)$, as a Lie algebra over $k$, with

$$a \cdot v = \rho_a(v)$$

for every $a \in sl_2(k)$ and $v \in V$.

Let us now take $k$ to be a field, so that $V$ is a vector space over $k$. If $\lambda \in k$, then put

$$V_{\lambda} = \{ v \in V : h \cdot v = \lambda v \},$$

305
which is the linear subspace of $V$ consisting of eigenvectors of $\rho_h$ with eigenvalue $\lambda$. Thus $V_\lambda \neq \{0\}$ exactly when $\lambda$ is an eigenvalue of $\rho_h$ on $V$. In this case, $\lambda$ may be called a *weight* of $h$ on $V$, and $V_\lambda$ may be called a *weight space*, as on p31 of [13]. An element of $V_\lambda$ may be said to have weight $\lambda$, as on p17 of [23].

Suppose that $v \in V_\lambda$ for some $\lambda \in k$. Observe that

\begin{align}
(15.1.6) \quad h \cdot (x \cdot v) &= ([h,x]) \cdot v + x \cdot (h \cdot v) \\
&= (2 \cdot x) \cdot v + x \cdot (\lambda v) = (\lambda + 2) (x \cdot v).
\end{align}

More precisely, $\lambda + 2$ means $\lambda + 2 \cdot 1 = \lambda + 1 + 1$ as an element of $k$, where $1$ is the multiplicative identity element in $k$. It follows that

\begin{equation}
(15.1.7) \quad x \cdot v \in V_{\lambda+2}.
\end{equation}

Similarly,

\begin{align}
(15.1.8) \quad h \cdot (y \cdot v) &= ([h,y]) \cdot v + y \cdot (h \cdot v) \\
&= (-2 \cdot y) \cdot v + y \cdot (\lambda v) = (\lambda - 2) (y \cdot v),
\end{align}

where $\lambda - 2 = \lambda - 2 \cdot 1 = \lambda - 1 - 1$, as an element of $k$. This means that

\begin{equation}
(15.1.9) \quad y \cdot v \in V_{\lambda-2}.
\end{equation}

Of course, if $k$ has characteristic 0, then $k$ may be considered as containing the rational numbers as a subfield.

Remember that nonzero eigenvectors of $\rho_h$ with distinct eigenvalues are automatically linearly independent. If $V$ has finite dimension as a vector space over $k$, then it follows that there are only finitely many weights $\lambda$ of $h$ on $V$.

### 15.2 Maximal or primitive vectors

Let us continue with the same notation and hypotheses as in the previous section. Suppose for the moment that $k$ is an algebraically closed field of characteristic 0, and that $V$ has positive finite dimension. This implies that $\rho_h$ is diagonalizable on $V$, as in Section 14.9. This also uses the fact that $h$ is ad-diagonalizable in $sl_2(k)$.

Under these conditions, there is a weight $\lambda \in k$ of $h$ on $V$, because $V \neq \{0\}$. In fact, there is a weight $\lambda$ of $h$ on $V$ such that $V_{\lambda+2} = \{0\}$, because there are only finitely many weights of $h$ on $V$. In this case, if $v \in V_\lambda$, then

\begin{equation}
(15.2.1) \quad x \cdot v = 0,
\end{equation}

by (15.1.7).

If $v$ is a nonzero element of $V_\lambda$ for some $\lambda \in k$ that satisfies (15.2.1), then $v$ may be called a *maximal vector of weight* $\lambda$, as on p32 of [13]. In this situation, $v$ may also be said to be *primitive of weight* $\lambda$, as in Definition 1 on p18 of [23]. Note that $V$ does not need to have finite dimension for this definition.
15.3. RELATED SUBMODULES

If \( k \) is an algebraically closed field of characteristic 0 and \( V \) has positive finite dimension, then one can get a maximal vector of some weight \( \lambda \in k \) by choosing \( \lambda \) so that \( V_{\lambda+2} = \{0\} \), as before. Alternatively, if \( v \in V \) is a nonzero eigenvector of \( p_h \) with eigenvalue \( \lambda \in k \), then

\[
(15.2.2) \quad p_h ((x^j)(v)) = (\lambda + 2j) (x^j)(v)
\]

for every nonnegative integer \( j \), as before. We also have that \( (x^j)(v) = 0 \) for some positive integer \( j \), because \( V \) has finite dimension. Let \( j_0 \) be the largest nonnegative integer such that \( (x^{j_0})(v) \neq 0 \), so that \( (x^{j_0+1})(v) = 0 \). This means that \( (x^{j_0})(v) \) is a maximal vector of weight \( \lambda + 2j_0 \), as in the alternate proof of Proposition 3 on p18 of [23].

Let \( B \) be the linear subspace of \( \mathfrak{sl}_2(k) \) spanned by \( x \) and \( h \). This is a Lie subalgebra of \( \mathfrak{sl}_2(k) \), which is solvable as a Lie algebra over \( k \). Let \( v \) be a nonzero element of \( V \), so that

\[
(15.2.3) \quad \{t v : t \in k\}
\]

is a one-dimensional linear subspace of \( V \). If \( v \) is primitive, then (15.2.3) is mapped into itself by \( p_h \) and \( p_x \). Equivalently, this means that (15.2.3) is a submodule of \( V \), as a module over \( B \), as a Lie algebra over \( k \).

Conversely, suppose that (15.2.3) is a submodule of \( V \), as a module over \( B \). This is the same as saying that \( v \) is an eigenvector of both \( p_h \) and \( p_x \). It follows that

\[
(15.2.4) \quad 2 : p_x(v) = ([p_h, p_x])(v) = 0.
\]

If \( k \) does not have characteristic 2, then we get that \( p_x(v) = 0 \), so that \( v \) is primitive. This corresponds to Proposition 2 on p18 of [23].

Suppose that \( k \) is an algebraically closed field of characteristic 0 again, and that \( V \) has positive finite dimension. Because \( B \) is solvable, there is a \( v \in V \) such that \( v \neq 0 \) and \( v \) is an eigenvector of \( p_h \) for every \( b \in B \), as in Section 9.13. This implies that \( v \) is primitive, as in the preceding paragraph. This corresponds to the proof of Proposition 3 on p18 of [23], and to Exercise 1 on p34 of [13].

15.3 Related submodules

Let \( k \) be a field of characteristic 0, and let \( V \) be a vector space over \( k \). Also let \( \rho = \rho^V \) be a representation of \( \mathfrak{sl}_2(k) \) on \( V \), as a Lie algebra over \( k \), so that \( V \) may be considered as a module over \( \mathfrak{sl}_2(k) \). Suppose that \( v \) is a primitive element of \( V \) of weight \( \lambda \in k \). Put

\[
(15.3.1) \quad v_j = (1/j!)(p_h)^j(v)
\]

for every positive integer \( j \), as on p32 of [13], and Theorem 3 on p18 of [23]. More precisely, \( (p_h)^j \) is the \( j \)th power of \( p_h \) with respect to composition, as a linear mapping from \( V \) into itself. We can define \( v_j \) as in (15.3.1) when \( j = 0 \) too, with the usual interpretations, so that \( v_0 = v \). It is convenient to put \( v_{-1} = 0 \) as well.
Under these conditions, we have that

\[(15.3.2) \quad h \cdot v_j = (\lambda - 2j) v_j\]

for every \(j \geq 0\), which is to say that \(v_j \in V_{\lambda - 2j}\). This follows by using (15.1.9) repeatedly, or using induction. We also have that

\[(15.3.3) \quad y \cdot v_j = (j + 1) \cdot v_{j+1}\]

for every \(j \geq -1\), by definition of \(v_j\). Let us check that

\[(15.3.4) \quad x \cdot v_j = (\lambda - j + 1) v_j - 1\]

for every \(j \geq 0\). If \(j = 0\), then this uses the hypothesis that \(v_0 = v\) be primitive, so that \(x \cdot v = 0\). Otherwise, if \(j \geq 1\), then we can use (15.3.3) to get that

\[(15.3.5) \quad j x \cdot v_j = x \cdot (y \cdot v_{j-1}) = ([x, y]) \cdot v_{j-1} + y \cdot (x \cdot v_{j-1})\]

Using induction, we may suppose that the analogue of (15.3.4) for \(j - 1\) holds, so that

\[(15.3.6) \quad j x \cdot v_j = h \cdot v_{j-1} + (\lambda - (j - 1) + 1) y \cdot v_{j-2}\]

It follows that

\[(15.3.7) \quad j x \cdot v_j = (\lambda - 2(j - 1)) v_{j-1} + (\lambda - j + 2) ((j - 2) + 1) v_{j-1},\]

by (15.3.2) and (15.3.3). This reduces to

\[(15.3.8) \quad j x \cdot v_j = ((\lambda - 2j + 2) + (\lambda - j + 2)(j - 1)) v_{j-1} = (\lambda + (\lambda - j)(j - 1)) v_{j-1} = j(\lambda - j + 1) v_{j-1}\]

This implies (15.3.4), as desired. This corresponds to the lemma on p32 of [13], and Theorem 3 on p18 of [23].

Remember that \(v_0 = v \neq 0\), because \(v\) is primitive, by hypothesis. As in Corollary 1 on p19 of [23], there are two cases to consider. In the first case,

\[(15.3.9) \quad \text{for each } j \geq 0, \text{ we have that } v_j \neq 0.\]

This means that the \(v_j\)'s are linearly independent in \(V\), because they are eigenvectors for \(P_0\) with distinct eigenvalues, by (15.3.2). In particular, this implies that \(V\) has infinite dimension as a vector space over \(k\).

In the second case,

\[(15.3.10) \quad v_l = 0 \text{ for some } l \geq 1.\]

This implies that \(v_j = 0\) for every \(j \geq l\), by the definition of \(v_j\). In this case, let \(m\) be the largest nonnegative integer such that \(v_m \neq 0\). Thus \(v_j \neq 0\) when \(j \leq m\), and \(v_j = 0\) when \(j > m\). Observe that

\[(15.3.11) \quad x \cdot v_{m+1} = (\lambda - m) v_m,\]
15.4. CONSTRUCTING MODULES $W(M)$

by (15.3.4) with $j = m + 1$. It follows that

$$\lambda = m,$$

because $v_m \neq 0$ and $v_{m+1} = 0$. This is also related to some of the remarks on p32 of [13].

In this second case, let $W$ be the linear span of $v_0, v_1, \ldots, v_m$ in $V$. It is easy to see that $W$ is a submodule of $V$, as a module over $sl_2(k)$, by (15.3.2), (15.3.3), and (15.3.4). Let $W_0$ be a nontrivial linear subspace of $W$ such that $\rho_h(W_0) \subseteq W_0$. Because $\rho_h$ is diagonalizable on $W$, we get that the restriction of $\rho_h$ to $W_0$ is diagonalizable as well. In particular, $W_0$ contains a nonzero eigenvector of $\rho_h$. In this situation, this means that $v_{j_0} \in W_0$ for some $j_0$, $0 \leq j_0 \leq m$, because the corresponding eigenvalues for $\rho_h$ on $W$ are distinct. If $W_0$ is a submodule of $W$, as a module over $sl_2(k)$, then it follows that $v_j \in W_0$ for every $j = 0, 1, \ldots, m$, by (15.3.3) and (15.3.4). This implies that $W_0 = W$, so that $W$ is irreducible as a module over $sl_2(k)$. This corresponds to Corollary 2 on p19 of [23], as well as some of the remarks on p32-3 and part of Exercise 3 on p34 of [13].

15.4 Constructing modules $W(m)$

Let $k$ be a commutative ring with a multiplicative identity element, and let $m$ be a nonnegative integer. Also let $W(m)$ be a module over $k$ freely generated by $m + 1$ nonzero distinct elements $v_0, v_1, \ldots, v_m$. This corresponds to $V(m)$ on p33 of [13], and to $W_m$ on p19 of [23]. More precisely, one can take $W(m)$ to be the space $k^{m+1}$ of $(m+1)$-tuples of elements of $k$, as a module over $k$ with coordinatewise addition and scalar multiplication, and $v_0, v_1, \ldots, v_m$ to be the standard basis vectors in $k^{m+1}$. It will be convenient to put $v_{-1} = v_{m+1} = 0$, as elements of $W(m)$.

Put

\begin{align*}
H(v_j) &= (m - 2j) \cdot v_j, \\
Y(v_j) &= (j + 1) \cdot v_{j+1}, \\
X(v_j) &= (m - j + 1) \cdot v_{j-1}
\end{align*}

for $j = 0, 1, \ldots, m$. More precisely, there are unique module homomorphisms $H$, $Y$, and $X$ from $W(m)$ into itself that satisfy these conditions. Note that (15.4.1) holds trivially when $j = -1$ or $m + 1$. Similarly, (15.4.2) holds trivially when $j = -1$, and (15.4.3) holds trivially when $j = m + 1$.

Observe that

\begin{align*}
H(X(v_j)) - X(H(v_j)) &= (m - j + 1) \cdot H(v_{j-1}) - (m - 2j) \cdot X(v_j) \\
&= (m - j + 1)(m - 2(j - 1)) \cdot v_{j-1} \\
&\quad - (m - 2j) \cdot X(v_j) \\
&= ((m - 2(j - 1)) - (m - 2j)) \cdot X(v_j) = 2 \cdot X(v_j)
\end{align*}
for each \( j = 0, 1, \ldots, m \). Similarly,

\[
H(Y(v_j)) - Y(H(v_j)) = (j + 1) \cdot H(v_{j+1}) - (m - 2j) \cdot Y(v_j)
\]

(15.4.5)

\[
= (j + 1)(m - 2(j + 1)) \cdot v_{j+1} - (m - 2j)Y(v_j)
\]

\[
= ((m - 2(j + 1)) - (m - 2j)) \cdot Y(v_j) = -2 \cdot Y(v_j)
\]

for every \( j = 0, 1, \cdots, m \). We also have that

\[
X(Y(v_j)) - Y(X(v_j)) = (j + 1) \cdot X(v_{j+1}) - (m - j + 1) \cdot Y(v_{j-1})
\]

(15.4.6)

\[
= (j + 1)(m - (j + 1) + 1) \cdot v_j
\]

\[
- (m - j + 1)(j - 1 + 1) \cdot v_j
\]

\[
= ((j + 1)(m - j) - (m - j + 1)) \cdot v_j
\]

\[
= ((m - j) - j) \cdot v_j = H(v_j)
\]

for each \( j = 0, 1, \ldots, n \). This shows that

\[
[H, X] = 2 \cdot X, \quad [H, Y] = -2 \cdot Y, \quad \text{and} \quad [X, Y] = H,
\]

using the commutator bracket associated to composition of module homomorphisms from \( W(m) \) into itself.

Let \( x, y, \) and \( h \) be the usual elements of \( sl_2(k) \), as in (15.1.1). Thus every element of \( sl_2(k) \) can be expressed in a unique way as a linear combination of \( x, y, \) and \( h \) with coefficients in \( k \), as in Section 10.2. The Lie brackets of \( x, y, \) and \( h \) in \( sl_2(k) \) satisfy (15.1.2), as before. Put

\[
(15.4.8) \quad \rho_x = X, \quad \rho_y = Y, \quad \text{and} \quad \rho_h = H.
\]

If \( a \in sl_2(k) \), then we can define \( \rho_a \) as a module homomorphism from \( W(m) \) into itself that satisfies (15.4.8) and is linear over \( k \) in \( a \). This defines a representation \( \rho = \rho^W(m) \) of \( sl_2(k) \), as a Lie algebra over \( k \), on \( W(m) \), because of (15.4.7). This corresponds to the statement before Theorem 2 on p20 of [23], as well as remarks on p33 and part of Exercise 3 on p34 of [13].

If \( m = 0 \), then \( \rho \) corresponds to the trivial representation of \( sl_2(k) \) on \( k \). If \( m = 1 \), then \( \rho \) corresponds to the standard representation of \( sl_2(k) \) on \( k^2 \). If \( m = 2 \), then one can check that \( \rho \) is isomorphic to the adjoint representation on \( sl_2(k) \). More precisely, if one takes \( v_0 = x, \ v_1 = -h, \) and \( v_2 = y \), then \( X, Y, \) and \( H \) correspond to \( \text{ad}_x, \text{ad}_y, \) and \( \text{ad}_h \) on \( sl_2(k) \), respectively. This corresponds to the examples mentioned on p20 of [23], and some remarks on p33 of [13].

### 15.5 More on \( W(m) \)

Let \( k \) be a field of characteristic 0, and let \( m \) be a nonnegative integer. Also let \( W(m) \) be as in the previous section, so that \( W(m) \) is a vector space over \( k \) of dimension \( m + 1 \), and \( \rho = \rho^W(m) \) is a representation of \( sl_2(k) \) on \( W(m) \). In this situation, \( v_0 \) is a primitive element of \( V = W(m) \) with weight \( \lambda = m \), by
construction. Of course, (15.3.2), (15.3.3), and (15.3.4) correspond to (15.4.1), (15.4.2), and (15.4.3) here, respectively. Thus \( v_1, \ldots, v_m \) in the previous section correspond to the analogous vectors in Section 15.3, and \( v_j = 0 \) when \( j > m \) in the notation of Section 15.3. This means that we are in the second case (15.3.10), and \( W(m) \) is the same as \( W \) in Section 15.3. In particular, \( W(m) \) is irreducible as a module over \( \mathfrak{sl}_2(k) \), as before. This corresponds to part (a) of Theorem 2 on p20 of [23], as well as remarks on p33 and part of Exercise 3 on p34 of [13].

Let \( V \) be a vector space over \( k \), and let \( p^V \) be a representation of \( \mathfrak{sl}_2(k) \) on \( V \). Suppose that \( v \in V \) is primitive of weight \( \lambda \in k \), as in Section 15.3, and that we are in the second case (15.3.10), which holds automatically when \( V \) has finite dimension. If \( W \subseteq V \) and \( m \geq 0 \) are as in Section 15.3, then \( W \) is isomorphic to \( W(m) \) in the previous section, as modules over \( \mathfrak{sl}_2(k) \). If \( V \) is irreducible as a module over \( \mathfrak{sl}_2(k) \), then \( V = W \), so that \( V \) is isomorphic to \( W(m) \) as a module over \( \mathfrak{sl}_2(k) \).

Suppose now that \( k \) is algebraically closed, and that \( V \) has positive finite dimension. Under these conditions, there is a \( v \in V \) that is primitive of some weight \( \lambda \in k \), as in Section 15.1. In this situation, we are automatically in the second case (15.3.10), as in the preceding paragraph. This corresponds to part (b) of Theorem 2 on p20 of [23], and the theorem on p33 of [13].

Let \( k \) be a field, and let \( W(m) \) be as in the previous section for each nonnegative integer \( m \) again. Note that \( W(0) \) is automatically irreducible as a module over \( \mathfrak{sl}_2(k) \). Similarly, \( W(1) \) is irreducible as a module over \( \mathfrak{sl}_2(k) \). Suppose from now on in this section that \( k \) has positive characteristic, and that \( m \geq 2 \).

Suppose for the moment that \( m \) is strictly less than the characteristic of \( k \), which is thus greater than 2. Under these conditions, \( W(m) \) is irreducible as a module over \( \mathfrak{sl}_2(k) \), as in Exercise 5 on p34 of [13]. To see this, let \( W_0 \) be a nontrivial linear subspace of \( W(m) \) such that \( H(W_0) \subseteq W_0 \). The restriction of \( H \) to \( W_0 \) is diagonalizable on \( W_0 \), because \( H \) is diagonalizable on \( W(m) \). This implies that \( W_0 \) contains a nonzero eigenvector of \( H \). One can check that the eigenvalues of \( H \) on \( W(m) \) are distinct as elements of \( k \) in this situation. It follows that \( v_{j_0} \in W_0 \) for some \( j_0, 0 \leq j_0 \leq m \).

Suppose that \( W_0 \) is a submodule of \( W(m) \), as a module over \( \mathfrak{sl}_2(k) \). One can verify that \( v_j \in W_0 \) for every \( j = 0, 1, \ldots, m \), using (15.4.2) and (15.4.3). More precisely, this uses the fact that the relevant coefficients on the right sides of (15.4.2) and (15.4.3) correspond to nonzero elements of \( k \). This implies that \( W_0 = W(m) \), as desired.

If \( k \) has characteristic equal to \( m \), then \( W(m) \) is reducible as a module over \( \mathfrak{sl}_2(k) \), as in Exercise 5 on p34 of [13]. To see this, consider the linear span \( W_1 \) of \( v_1, \ldots, v_{m-1} \) in \( W(m) \). Clearly \( H(W_1) \subseteq W_1 \), because the \( v_i \)'s are eigenvectors for \( H \). It is easy to see that \( Y(W_1) \subseteq W_1 \), because \( Y(v_{m-1}) = 0 \) when the characteristic of \( k \) is equal to \( m \), by (15.4.2). Similarly, one can check that \( X(W_1) \subseteq W_1 \), because \( X(v_1) = 0 \) when the characteristic of \( k \) is \( m \), by (15.4.3). This implies that \( W_1 \) is a submodule of \( W(m) \), as a module over \( \mathfrak{sl}_2(k) \). Of course, \( W_1 \neq \{0\}, W(m) \).
15.6 Another construction

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( T_1, T_2 \) be commuting indeterminates. Remember that \( k[T_1, T_2] \) is the space of formal polynomials in \( T_1, T_2 \) with coefficients in \( k \), as in Section 5.8. This is a commutative associative algebra over \( k \) with respect to the usual formal multiplication of polynomials, as before. If \( d \) is a nonnegative integer, then the space \( k_d[T_1, T_2] \) of formal polynomials in \( T_1, T_2 \) with coefficients in \( k \) that are homogeneous of degree \( d \) is a submodule of \( k[T_1, T_2] \), as a module over \( k \), as in Section 5.13. More precisely, \( k[T_1, T_2] \) corresponds to the direct sum of the submodules \( k_d[T_1, T_2] \) over \( d \geq 0 \), as a module over \( k \), as before.

We would like to make \( k[T_1, T_2] \) into a module over \( sl_2(k) \), as a Lie algebra over \( k \), as in Exercise 4 on p34 of [13], and the remark on p20 of [23]. Thus, if \( a \in sl_2(k) \) and \( f(T) \in k[T_1, T_2] \), then we need to define \( a \cdot f(T) \) as an element of \( k[T_1, T_2] \). More precisely, if \( f(T) \in k_d[T_1, T_2] \) for some \( d \geq 0 \), then \( a \cdot f(T) \) will be an element of \( k_d[T_1, T_2] \). This means that \( k_d[T_1, T_2] \) will be a submodule of \( k[T_1, T_2] \), as a module over \( sl_2(k) \), for each \( d \geq 0 \). Indeed, \( k[T_1, T_2] \) will correspond to the direct sum of \( k_d[T_1, T_2] \) over \( d \geq 0 \), as a module over \( sl_2(k) \).

Of course, \( k_0[T_1, T_2] \) corresponds to constant formal polynomials with coefficients in \( k \), and is isomorphic to \( k \) in an obvious way. If \( a \in sl_2(k) \) and \( f(T) \in k_0[T_1, T_2] \), then we put \( a \cdot f(T) = 0 \).

An element of \( k_1[T_1, T_2] \) can be expressed in a unique way as

\[
(15.6.1) \quad f(T) = f_1 T_1 + f_2 T_2,
\]

where \( f_1, f_2 \in k \). In the notation of Section 5.8, \( f_1 \) and \( f_2 \) correspond to \( f_\alpha \) with \( \alpha = (1, 0) \) and \( (0, 1) \), respectively. If \( a \in sl_2(k) \), then we can define \( a \cdot f(T) \) as an element of \( k_1[T_1, T_2] \) using the standard action of \( sl_2(k) \) on \( (f_1, f_2) \) as an element \( k^2 \). In particular, if \( x, y, h \in sl_2(k) \) are as in (15.1.1), as usual, then

\[
(15.6.2) \quad x \cdot f(T) = f_2 T_1, \quad y \cdot f(T) = f_1 T_2, \quad h \cdot f(T) = f_1 T_1 - f_2 T_2.
\]

This makes \( k_1[T_1, T_2] \) into a module over \( sl_2(k) \), as a Lie algebra over \( k \).

If \( a \in sl_2(k) \), then we would like to define the action of \( a \) on \( k_1[T_1, T_2] \) so that it is a derivation, as in Exercise 4 on p34 of [13]. Thus we would like to have that

\[
(15.6.3) \quad a \cdot (f(T) g(T)) = (a \cdot f(T)) g(T) + f(T) (a \cdot g(T))
\]

for every \( f(T), g(T) \in k[T_1, T_2] \). If \( \alpha = (\alpha_1, \alpha_2) \) is an ordered pair of nonnegative integers, then we should have that

\[
(15.6.4) \quad a \cdot (T_1^{\alpha_1} T_2^{\alpha_2}) = \alpha_1 \cdot (a \cdot T_1) T_1^{\alpha_1 - 1} T_2^{\alpha_2} + \alpha_2 \cdot (a \cdot T_2) T_1^{\alpha_1 - 1} T_2^{\alpha_2}.\]

More precisely, if \( \alpha_1 \) or \( \alpha_2 \) is 0, then the corresponding term on the right should be interpreted as being equal to 0. One can use this to define the left side of the equation, using the definition of \( a \cdot T_1 \) and \( a \cdot T_2 \) from the preceding paragraph. This can be used to define \( a \cdot f(T) \) for every \( f(T) \in k[T_1, T_2] \), by linearity. Using this definition, one can check that (15.6.3) holds for every \( f(T), g(T) \in k[T_1, T_2] \).
Let \( \rho_a \) be the action of \( a \in \mathfrak{sl}_2(k) \) on \( k[T_1, T_2] \) defined in the previous paragraph. Thus \( \rho_a \) is a derivation on \( k[T_1, T_2] \) for every \( a \in \mathfrak{sl}_2(k) \), and it is easy to see that \( \rho_a \) is linear over \( k \) in \( a \). We also have that \( \rho_a \) maps \( k_0[T_1, T_2] \) into itself for every \( a \in \mathfrak{sl}_2(k) \) and \( d \geq 0 \), by construction. In order to show that this defines a representation of \( \mathfrak{sl}_2(k) \), as a Lie algebra over \( k \), on \( k[T_1, T_2] \), we should verify that

\[
[\rho_a, \rho_b] = \rho_{[a,b]}
\]

for every \( a, b \in \mathfrak{sl}_2(k) \). Note that the left side is a derivation on \( k[T_1, T_2] \) as well, as in Section 2.5. We already have that \( \rho \) defines a representation of \( \mathfrak{sl}_2(k) \) on \( k_0[T_1, T_2] \) and \( k_1[T_1, T_2] \), so that (15.6.5) holds on these subspaces of \( k[T_1, T_2] \). One can use this to check that (15.6.5) holds on all of \( k[T_1, T_2] \), because both sides of (15.6.5) are derivations on \( k[T_1, T_2] \).

Let \( \alpha = (\alpha_1, \alpha_2) \) be an ordered pair of nonnegative integers again. Observe that

\[
\begin{align*}
(15.6.6) & \quad x \cdot (T_1^{\alpha_1} T_2^{\alpha_2}) = \alpha_2 \cdot T_1^{\alpha_1+1} T_2^{\alpha_2-1}, \\
(15.6.7) & \quad y \cdot (T_1^{\alpha_1} T_2^{\alpha_2}) = \alpha_1 \cdot T_1^{\alpha_1-1} T_2^{\alpha_2+1}, \\
(15.6.8) & \quad h \cdot (T_1^{\alpha_1} T_2^{\alpha_2}) = (\alpha_1 - \alpha_2) \cdot T_1^{\alpha_1} T_2^{\alpha_2}.
\end{align*}
\]

If \( a \in \mathfrak{sl}_2(k) \), then \( \rho_a \) corresponds to the formal differential operator

\[
(15.6.9) \quad (a \cdot T_1) \frac{\partial}{\partial T_1} + (a \cdot T_2) \frac{\partial}{\partial T_2} = (a \cdot T_1) \frac{\partial}{\partial T_1} + (a \cdot T_2) \frac{\partial}{\partial T_2}
\]

on \( k[T_1, T_2] \), in the notation of Section 5.11. More precisely, this is a first-order differential operator, as in Section 5.12, which is homogeneous of degree 0, as in Section 5.14.

### 15.7 Some exponentials

Let \( k \) be a field of characteristic 0, or a commutative ring with a multiplicative identity element that is an algebra over \( \mathbb{Q} \), for simplicity. As in (15.1.1), we put \( x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \) and \( y = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \). Note that \( x^2 = y^2 = 0 \), so that \( x \) and \( y \) are nilpotent as elements of the algebra \( M_2(k) \) of \( 2 \times 2 \) matrices with entries in \( k \). Thus the exponentials of \( x \) and \( y \) are defined in \( M_2(k) \) as in Section 14.10, with

\[
\begin{align*}
(15.7.1) & \quad \exp x = I + x = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \\
& \quad \exp y = I + y = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},
\end{align*}
\]

where \( I = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \) is the identity matrix in \( M_2(k) \). Similarly,

\[
\begin{align*}
(15.7.2) & \quad \exp(-y) = I - y = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.
\end{align*}
\]

Observe that

\[
(15.7.3) \quad (\exp x) (\exp -y) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}.
\]
and hence

\[(15.7.4) \quad (\exp x)(\exp -y)(\exp x) = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.\]

Remember that \([x, y] = h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, [h, x] = 2 \cdot x, \text{ and } [h, y] = -2 \cdot y.\]

Put \(\text{ad}_x(z) = [x, z]\) and \(\text{ad}_y(z) = [y, z]\) for every \(z \in \text{sl}_2(k)\), as usual. Thus \(\text{ad}_x(x) = \text{ad}_y(y) = 0,\)

\[(15.7.5) \quad \text{ad}_x(y) = -\text{ad}_y(x) = h, \quad \text{ad}_x(h) = -2 \cdot x, \quad \text{ad}_y(h) = 2 \cdot y.\]

This implies that

\[(15.7.6) \quad (\text{ad}_x)^2(h) = (\text{ad}_y)^2(h) = 0.\]

Similarly,

\[(15.7.7) \quad (\text{ad}_x)^2(y) = \text{ad}_x(h) = -2 \cdot x, \quad (\text{ad}_y)^2(x) = \text{ad}_y(-h) = -2 \cdot y.\]

It follows that \((\text{ad}_x)^3 = (\text{ad}_y)^3 = 0\) as mappings from \(\text{sl}_2(k)\) into itself, so that \(x\) and \(y\) are ad-nilpotent in \(\text{sl}_2(k)\).

This means that the exponentials of \(\text{ad}_x\) and \(\text{ad}_y\) are defined as linear mappings from \(\text{sl}_2(k)\) into itself, as in Section 14.10. In this situation, we have that

\[(15.7.8) \quad \exp \text{ad}_x = I + \text{ad}_x + (1/2) (\text{ad}_x)^2,\]

\[(15.7.9) \quad \exp \text{ad}_y = I + \text{ad}_y + (1/2) (\text{ad}_y)^2,\]

where \(I\) is the identity mapping on \(\text{sl}_2(k)\). Similarly,

\[(15.7.10) \quad \exp(-\text{ad}_y) = I - \text{ad}_y + (1/2) (\text{ad}_y)^2.\]

Observe that

\[(15.7.11) \quad (\exp \text{ad}_x)(x) = x, \quad (\exp \text{ad}_x)(h) = h - 2 \cdot x, \quad (\exp \text{ad}_x)(y) = y + h - x.\]

We also have that

\[(15.7.12) \quad (\exp -\text{ad}_y)(y) = y, \quad (\exp -\text{ad}_y)(h) = h - 2 \cdot y, \quad (\exp -\text{ad}_y)(x) = x + h - y.\]

Put

\[(15.7.13) \quad \sigma = (\exp \text{ad}_x) \circ (\exp -\text{ad}_y) \circ (\exp \text{ad}_x),\]

which defines a linear mapping from \(\text{sl}_2(k)\) into itself. More precisely, this is an inner automorphism of \(\text{sl}_2(k)\), as in Section 14.11. Observe that

\[(15.7.14) \quad \sigma(x) = (\exp \text{ad}_x)(x + h - y) = x + h - 2 \cdot x - y - h + x = -y.\]
15.8. Exponentials and \( W(M) \)

Similarly,

\[
\sigma(y) = (\exp \text{ad}_x)((\exp - \text{ad}_y)(y + h - x))
\]

(15.7.15) \[= (\exp \text{ad}_x)(y + h - 2 \cdot y - x + y) = (\exp \text{ad}_x)(-x) = -x. \]

We also have that

\[
\sigma(h) = (\exp \text{ad}_x)((\exp - \text{ad}_y)(h - 2 \cdot x))
\]

(15.7.16) \[= (\exp \text{ad}_x)(h - 2 \cdot y - 2 \cdot x + 2 \cdot y + h) = (\exp \text{ad}_x)(-h + 2 \cdot x - 2 \cdot x = -h. \]

It follows from (15.7.14), (15.7.15), and (15.7.16) that

\[
\sigma(z) = -z^t
\]

(15.7.17) for every \( z \in \mathfrak{sl}_2(k) \), where \( z^t \) is the transpose of \( z \). This is the same as conjugating \( z \in \mathfrak{sl}_2(k) \) by (15.7.4), as in Section 14.11. This corresponds to some remarks on p9 and Exercise 10 on p10 of [13].

15.8 Exponentials and \( W(m) \)

Let \( k \) be a commutative ring with a multiplicative identity element that is an algebra over \( \mathbb{Q} \), or simply a field of characteristic 0, and let \( m \) be a nonnegative integer. Remember that \( W(m) \) is defined as a module over \( k \) as in Section 15.4, with corresponding linear mappings \( X, Y, \) and \( H \). Note that \( X \) and \( Y \) are nilpotent as linear mappings on \( W(m) \), so that their exponentials are defined as invertible linear mappings on \( W(m) \), as in Section 14.10. Put

\[
(15.8.1) \quad \theta = (\exp X) \circ (\exp - Y) \circ (\exp X),
\]

which is an invertible linear mapping from \( W(m) \) into itself.

Let \( \rho \) be the representation of \( \mathfrak{sl}_2(k) \) on \( W(m) \) corresponding to \( X, Y, \) and \( H \) as before. Thus \( \rho_x = X, \rho_y = Y, \) and \( \rho_h = H \), where \( x, y, \) and \( h \) are the usual elements of \( \mathfrak{sl}_2(k) \). Remember that \( \exp \text{ad}_x \) and \( \exp \text{ad}_y \) are defined as invertible linear mappings on \( \mathfrak{sl}_2(k) \), because \( x \) and \( y \) are ad-nilpotent in \( \mathfrak{sl}_2(k) \).

If \( z \in \mathfrak{sl}_2(k) \), then

\[
(15.8.2) \quad \rho_{(\exp \text{ad}_x)(z)} = (\exp \rho_x) \circ \rho_z \circ (\exp \rho_x)^{-1},
\]

as in Section 14.12. Similarly,

\[
(15.8.3) \quad \rho_{(\exp - \text{ad}_y)(z)} = (\exp - \rho_y) \circ \rho_z \circ (\exp - \rho_y)^{-1}
\]

for every \( z \in \mathfrak{sl}_2(k) \).

If \( \sigma \) is as in (15.7.13), then we get that

\[
(15.8.4) \quad \rho_{\sigma(z)} = \theta \circ \rho_z \circ \theta^{-1}
\]
for every \( z \in \mathfrak{sl}_2(k) \). In particular,

\[(15.8.5)\quad -Y = \rho_{\sigma(z)} = \theta \circ \rho_z \circ \theta^{-1} = \theta \circ X \circ \theta^{-1},\]

using (15.7.14) in the first step. Similarly,

\[(15.8.6)\quad -X = \rho_{\sigma(y)} = \theta \circ \rho_y \circ \theta^{-1} = \theta \circ Y \circ \theta^{-1},\]

using (15.7.15) in the first step. We also have that

\[(15.8.7)\quad -H = \rho_{\sigma(h)} = \theta \circ \rho_h \circ \theta^{-1} = \theta \circ H \circ \theta^{-1},\]

using (15.7.16) in the first step. This corresponds to remarks on p334 of [13], and Remark 1 on p21 of [23].

### 15.9 Additional properties of \( \mathfrak{sl}_2(k) \) modules

Let \( K \) be a field, and let \( k \) be a subfield of \( K \). Also let \( n \) be a positive integer, so that \( V = k^n \) is an \( n \)-dimensional vector space over \( k \), and \( V_K = K^n \) is an \( n \)-dimensional vector space over \( K \). If \( T \) is a linear mapping from \( V \) into itself, as a vector space over \( k \), then there is a unique extension \( T_K \) of \( T \) to a linear mapping from \( V_K \) into itself, as a vector space over \( K \). If \( T \) is invertible on \( V \), then \( T_K \) is invertible on \( V_K \), with inverse equal to the extension of \( T^{-1} \) on \( V \) to \( V_K \) as before. If \( \lambda \in k \) is an eigenvalue of \( T \) on \( V \), then \( \lambda \) is an eigenvalue of \( T_K \) on \( V_K \). However, if \( \lambda \in k \) is not an eigenvalue of \( T \) on \( V \), then \( \lambda \) is not an eigenvalue of \( T_K \) on \( V_K \). More precisely, \( T - \lambda I_V \) is invertible on \( V \) in this case, where \( I_V \) is the identity mapping on \( V \). This implies that \( T_K - \lambda I_{V_K} \) is invertible on \( V_K \), where \( I_{V_K} \) is the identity mapping on \( V_K \). Thus the eigenvalues of \( T \) on \( V \) are the same as the eigenvalues of \( T_K \) on \( V_K \) that are elements of \( k \).

Let \( \rho^V \) be a representation of \( \mathfrak{sl}_2(k) \), as a Lie algebra over \( k \), on \( V = k^n \). This leads to a representation \( \rho^{V_K} \) of \( \mathfrak{sl}_2(K) \), as a Lie algebra over \( K \), on \( V_K = K^n \), in a natural way. More precisely, if \( z \in \mathfrak{sl}_2(k) \), then \( \rho^{V_K}_z \) is the \( K \)-linear mapping on \( K^n \) that corresponds to \( \rho^V_z \) as a \( k \)-linear mapping on \( k^n \) in the preceding paragraph. In particular, this can be applied to the usual elements \( x, y, \) and \( h \) of \( \mathfrak{sl}_2(k) \), as in (15.1.1). If \( z \in \mathfrak{sl}_2(K) \), then \( z \) can be expressed as a linear combination of \( x, y, \) and \( h \) with coefficients in \( K \), and one can take \( \rho^{V_K}_z \) to be the corresponding linear combination of \( \rho^{V_K}_x, \rho^{V_K}_y, \) and \( \rho^{V_K}_h \).

Suppose now that \( k \) has characteristic 0, and that \( K \) is algebraically closed. Of course, we may consider \( Q \) as a subfield of \( k \), and thus as a subfield of \( K \). Remember that \( V_K \) can be expressed as the direct sum of finitely many irreducible \( \mathfrak{sl}_2(K) \) modules, by Weyl’s theorem. Each of these irreducible \( \mathfrak{sl}_2(K) \) modules is isomorphic to the analogue of \( W(m) \) in Section 15.4 for \( K \), and for some nonnegative integer \( m \), as in Section 15.5. It follows that \( \rho^{V_K}_h \) is diagonalizable on \( V_K \), with eigenvalues in \( \mathbb{Z} \).

In particular, \( \rho^{V_K}_h \) has an eigenvalue \( \lambda \) in \( \mathbb{Z} \), which implies that \( \lambda \) is an eigenvalue for \( \rho^V_h \) on \( V \), as before. One can use this to show that there is a maximal or primitive vector in \( V \) of some weight, as in Section 15.2. If \( V \) is
irreducible as an $sl_2(k)$ module, then it follows that $V$ is isomorphic as an $sl_2(k)$ module to $W(m)$ for some nonnegative integer $m$, as in Section 15.5. Otherwise, one can use Weyl’s theorem to reduce to this case. This corresponds to Exercise 2 on p62 of [24], with $n = 2$.

15.10 Complexifying real vector spaces

Let $V_1$ be a vector space over the real numbers. The complexification of $V_1$ may be defined as a vector space $V_2$ over the complex numbers as follows. We start by taking $V_2$ to be the direct sum $V_1 \oplus V_1$ of $V_1$ with itself, as a vector space over the real numbers. Equivalently, $V_2$ is the Cartesian product $V_1 \times V_1$ of $V_1$ with itself, where addition and scalar multiplication by real numbers is defined coordinatewise. If $(v, w) \in V_1 \times V_1$, then $i(v, w)$ is defined as an element of $V_1 \times V_1$ by

\[ i(v, w) = (-w, v). \]

(15.10.1)

It is easy to see that this makes $V_2$ into a vector space over the complex numbers. We shall normally identify $v \in V_1$ with $(v, 0) \in V_1 \times V_1 = V_2$, so that $V_1$ may be considered as a subset of $V_2$. Note that any basis for $V_1$ as a vector space over the real numbers may be considered as a basis for $V_2$ as a vector space over the complex numbers.

Let $V$ be a vector space over the complex numbers, so that $V$ may be considered as a vector space over the real numbers too. Suppose that $V_0$ is a real-linear subspace of $V$, which is to say that $V_0$ is a linear subspace of $V$ as a vector space over the real numbers. In this case,

\[ iV_0 = \{ i \ v : v \in V_0 \} \]

(15.10.2)

is a real-linear subspace of $V$ as well. Note that $V_0 + iV_0$ is a complex-linear subspace of $V$, which is to say a linear subspace of $V$ as a vector space over the complex numbers. Let us say that $V_0$ is totally real in $V$ if

\[ V_0 \cap (iV_0) = \{ 0 \}. \]

(15.10.3)

This means that $V_0 + iV_0$ is isomorphic to the direct sum of $V_0$ and $iV_0$, as a vector space over the real numbers. Under these conditions, there is a natural isomorphism from the complexification of $V_0$ onto $V_0 + iV_0$, as vector spaces over the complex numbers, which is the identity mapping on $V_0$.

Let $V_1$ be a vector space over the real numbers again, and let $W$ be a vector space over the complex numbers. Thus $W$ may be considered as a vector space over the real numbers, and a linear mapping from $V_1$ into $W$, as a vector space over $\mathbf{R}$, may be called a real-linear mapping. One can verify that a real-linear mapping from $V_1$ into $W$ has a unique extension to a complex-linear mapping from the complexification $V_2$ of $V_1$ into $W$.

Let $W_1$ be a vector space over the real numbers, and let $W_2$ be the complexification of $W_1$. If $\phi_1$ is a linear mapping from $V_1$ into $W_1$, as vector spaces over the real numbers, then $\phi_1$ may be considered as a real-linear mapping from $V_1$ into $W_1$.
into $W_2$, because $W_1$ is identified with a real-linear subspace of $W_2$. Thus $\phi_2$
has a unique extension to a complex-linear mapping $\phi_2$ from $V_2$ into $W_2$, as in
the preceding paragraph. If $\phi_1$ is injective as a mapping from $V_1$ into $W_1$, then
one can check that $\phi_2$ is injective as a mapping from $V_2$ into $W_2$. Similarly, if
$\phi_1$ maps $V_1$ onto $W_1$, then $\phi_2$ maps $V_2$ onto $W_2$.

Let $W$ be a vector space over the complex numbers again, let $\phi_1$ be a real-
linear mapping from $V_1$ into $W$, and let $\phi_2$ be the extension of $\phi_1$ to a complex-
linear mapping from $V_2$ into $W$. If $\phi_1$ is injective as a mapping from $V_1$ into $W$,
and if $\phi_1(V_1)$ is totally real as a real-linear subspace of $W$, then one can verify
that $\phi_2$ is injective as a mapping from $V_2$ into $W$.

Let $W_1$ be a vector space over the real numbers again, and let $Z_1$ be another
vector space over the real numbers. Also let $\phi_1$ be a linear mapping from $V_1$
into $W_1$, and let $\psi_1$ be a linear mapping from $W_1$ into $Z_1$, as vector spaces over
$R$. As before, $\phi_1$ and $\psi_1$ have unique complex-linear extensions $\phi_2$ from $V_2$
into $W_2$ and $\psi_2$ from $W_2$ into the complexification $Z_2$ of $Z_1$, respectively. Of course,$\psi_1 \circ \phi_1$ is a linear mapping from $V_1$ into $Z_1$, as vector spaces over $R$. Note that
$\psi_2 \circ \phi_2$ is the unique extension of $\psi_1 \circ \phi_1$ to a complex-linear mapping from $V_2$
into $Z_2$.

Let $Z$ be a vector space over the complex numbers again. A mapping from
$V_1 \times W_1$ into $Z$ that is bilinear over $R$, where $Z$ is considered as a vector space
over the real numbers, may be called a real-bilinear mapping. One can check
that such a mapping has a unique extension to a mapping from $V_2 \times W_2$ into $Z$
that is bilinear over $C$. If $Z_1$ is a vector space over the real numbers, then
a mapping from $V_1 \times W_1$ into $Z_1$ that is bilinear over $R$ may be considered as
a real-bilinear mapping from $V_1 \times W_1$ into the complexification $Z_2$ of $Z_1$. This
can be extended to a mapping from $V_2 \times W_2$ into $Z_2$ that is bilinear over $C$, as
before.

\section{Spaces of linear mappings}

Let $V_1$ be a vector space over the real numbers, and let $V_2$ be its complexification,
as in the previous section. If $W$ is a vector space over the complex numbers,
then we let $\mathcal{L}_R(V_1, W)$ be the space of real-linear mappings from $V_1$ into $W$,
and $\mathcal{L}_C(V_2, W)$ be the space of complex-linear mappings from $V_2$ into $W$. If
$\phi_1$ is a real-linear mapping from $V_1$ into $W$, then there is a unique extension
$\phi_2$ of $\phi_1$ to a complex-linear mapping from $V_2$ into $W$, as before. Conversely,
if $\phi_2$ is any complex-linear mapping from $V_2$ into $W$, then the restriction of $\phi_2$
to $V_1$, considered as a real-linear subspace of $V_2$, is a real-linear mapping from
$V_1$ into $W$. This defines a one-to-one correspondence between $\mathcal{L}_R(V_1, W)$ and
$\mathcal{L}_C(V_2, W)$.

Note that $\mathcal{L}_R(V_1, W)$ may be considered as a vector space over the complex
numbers, with respect to pointwise addition and scalar multiplication by com-
plex numbers of mappings into $W$. It is easy to see that this corresponds to pointwise addition and scalar multiplication by complex numbers on $\mathcal{L}_C(V_2, W)$.
Thus the one-to-one correspondence between $\mathcal{L}_R(V_1, W)$ and $\mathcal{L}_C(V_2, W)$ men-
tioned in the preceding paragraph is an isomorphism between complex vector spaces.

Let \( W_1 \) be another vector over the real numbers, and let \( W_2 \) be its complexification, as before. The space \( \mathcal{L}(V_1, W_1) \) of linear mappings from \( V_1 \) into \( W_1 \) may be considered as a real-linear subspace of the space \( \mathcal{L}(V_1, W_2) \) of real-linear mappings from \( V_1 \) into \( W_2 \), because \( W_1 \) is identified with a real-linear subspace of \( W_2 \). Similarly, \( iW_1 \) may be considered as a real-linear subspace of \( W_2 \), so that the space \( \mathcal{L}(V_1, iW_1) \) of linear mappings from \( V_1 \) into \( iW_1 \), as vector spaces over \( \mathbb{R} \), may be considered as a real-linear subspace of \( \mathcal{L}(V_1, W_2) \). Observe that

\[
\mathcal{L}(V_1, iW_1) = i\mathcal{L}(V_1, W_1),
\]

as real-linear subspaces of \( \mathcal{L}(V_1, W_2) \). It is easy to see that

\[
\mathcal{L}(V_1, W_1) \cap (i\mathcal{L}(V_1, W_1)) = \{0\},
\]

because \( W_1 \cap (iW_1) = \{0\} \) in \( W_2 \). We also have that

\[
\mathcal{L}(V_1, W_1) + \mathcal{L}(V_1, iW_1) = \mathcal{L}(V_1, W_2).
\]

More precisely, \( \mathcal{L}(V_1, W_2) \) corresponds to the direct sum of \( \mathcal{L}(V_1, W_1) \) and \( i\mathcal{L}(V_1, W_1) \) as a vector space over \( \mathbb{R} \), because \( W_2 \) corresponds to the direct sum of \( W_1 \) and \( iW_1 \) as a vector space over \( \mathbb{R} \). Equivalently, this means that

\[
\mathcal{L}(V_1, W_1) + i\mathcal{L}(V_1, W_1) = \mathcal{L}(V_1, W_2).
\]

Thus the complexification of \( \mathcal{L}(V_1, W_2) \) can be identified with \( \mathcal{L}(V_1, W_2) \), as a vector space over the complex numbers, as in the previous section.

Remember that there is a natural isomorphism between \( \mathcal{L}(V_1, W_2) \) and \( \mathcal{L}(V_2, W_2) \), as before. This leads to a natural embedding of \( \mathcal{L}(V_1, W_2) \) into \( \mathcal{L}(V_2, W_2) \). More precisely, this embedding takes a linear mapping \( \phi_1 \) from \( V_1 \) into \( W_1 \), and associates to it the unique extension \( \phi_2 \) of \( \phi_1 \) to a complex-linear mapping from \( V_2 \) into \( W_2 \), where \( V_1, W_1 \) are identified with real-linear subspaces of \( V_2, W_2 \), as usual. The image of \( \mathcal{L}(V_1, W_1) \) in \( \mathcal{L}(V_2, W_2) \) consists of the complex-linear mappings \( \phi_2 \) from \( V_2 \) into \( W_2 \) such that

\[
\phi_2(V_1) \subseteq W_1.
\]

Using this embedding, we get an isomorphism between the complexification of \( \mathcal{L}(V_1, W_1) \) and \( \mathcal{L}(V_2, W_2) \).

Let \( n \) be a positive integer, so that \( \mathbb{R}^n \) is a vector space over the real numbers, whose complexification can be identified with \( \mathbb{C}^n \). Remember that the space \( \mathcal{L}(\mathbb{R}^n) \) of linear mappings from \( \mathbb{R}^n \) into itself is isomorphic to the space \( M_n(\mathbb{R}) \) of \( n \times n \) matrices with entries in \( \mathbb{R} \) in the usual way. Similarly, the space \( \mathcal{L}(\mathbb{C}^n) \) of complex-linear mappings from \( \mathbb{C}^n \) into itself is isomorphic to the space \( M_n(\mathbb{C}) \) of \( n \times n \) matrices with entries in \( \mathbb{C} \) in essentially the same way. Of course, \( M_n(\mathbb{R}) \) may be considered as a real-linear subspace of \( M_n(\mathbb{C}) \), and \( M_n(\mathbb{C}) \) can be identified with the complexification of \( M_n(\mathbb{R}) \). This corresponds to identifying a linear mapping from \( \mathbb{R}^n \) into itself with a complex-linear mapping from \( \mathbb{C}^n \) into itself that takes \( \mathbb{R}^n \) into itself, as in the preceding paragraph.
15.12 Complexifying algebras over $\mathbb{R}$

Let $A_1$ be an algebra over the real numbers in the strict sense, and let $A_2$ be the complexification of $A_1$ as a vector space over $\mathbb{R}$, as in Section 15.10. The algebra structure on $A_1$ corresponds to a mapping from $A_1 \times A_1$ into $A_1$ that is bilinear over $\mathbb{R}$, which has a unique extension to a mapping from $A_2 \times A_2$ into $A_2$ that is bilinear over $\mathbb{C}$, as before. This makes $A_2$ into an algebra over $\mathbb{C}$ in the strict sense. If $A_1$ is commutative, associative, or a Lie algebra, then one can check that $A_2$ has the analogous property, as an algebra over $\mathbb{C}$. Similarly, if $A_1$ has a multiplicative identity element $e$, then $e$ is the multiplicative identity element in $A_2$ as well.

Let $B$ be an algebra over the complex numbers in the strict sense, which may be considered as an algebra over the real numbers in the strict sense too. If $\phi_1$ is a real-linear mapping from $A_1$ into $B$, then there is a unique extension of $\phi_1$ to a complex-linear mapping $\phi_2$ from $A_2$ into $B$, as in Section 15.10. If $\phi_1$ is an algebra homomorphism from $A_1$ into $B$, then it is easy to see that $\phi_2$ is an algebra homomorphism from $A_2$ into $B$. Similarly, if $\phi_1$ is an opposite algebra homomorphism from $A_1$ into $B$, then $\phi_2$ is an opposite algebra homomorphism from $A_2$ into $B$.

If $B_1$ is an algebra over the real numbers in the strict sense, then its complexification $B_2$ is an algebra over the complex numbers in the strict sense, as before. If $\phi_1$ is a linear mapping from $A_1$ into $B_1$, as vector spaces over the real numbers, then $\phi_1$ may be considered as a real-linear mapping from $A_1$ into $B_2$. Thus there is a unique extension of $\phi_1$ to a complex-linear mapping $\phi_2$ from $A_2$ into $B_2$, as before. If $\phi_1$ is an algebra homomorphism or an opposite algebra homomorphism from $A_1$ into $B_1$, then $\phi_2$ has the analogous property as a mapping from $A_2$ into $B_2$.

Let $A$ be an algebra over the complex numbers in the strict sense, which may also be considered as an algebra over the real numbers in the strict sense. Suppose that $A_0$ is a real-linear subspace of $A$ that is totally real in $A$. This implies that there is a natural isomorphism between the complexification of $A_0$, as a vector space over $\mathbb{R}$, and $A_0 + iA_0$, as in Section 15.10. If $A_0$ is a subalgebra of $A$ too, as an algebra over $\mathbb{R}$ in the strict sense, then it is easy to see that $A_0 + iA_0$ is a subalgebra of $A$, as an algebra over $\mathbb{C}$ in the strict sense. In this case, $A_0 + iA_0$ is isomorphic to the complexification of $A_0$, as an algebra over $\mathbb{R}$ in the strict sense.

Let $V_1$ be a vector space over the real numbers, with complexification $V_2$, as in Section 15.10. Remember that the space $\mathcal{L}(V_1)$ of linear mappings from $V_1$ into itself is an associative algebra over $\mathbb{R}$ with respect to composition of mappings. Similarly, the space $\mathcal{L}_\mathbb{C}(V_2)$ of complex-linear mappings from $V_2$ into itself is an associative algebra over $\mathbb{C}$ with respect to composition of mappings. We have also seen that $\mathcal{L}_\mathbb{C}(V_2)$ can be identified with the complexification of $\mathcal{L}(V_1)$, as a vector space over $\mathbb{R}$, as in the previous section. More precisely, this uses the correspondence between linear mappings from $V_1$ into itself and complex-linear mappings from $V_2$ into itself that map $V_1$ into itself, where $V_1$ is considered as a real-linear subspace of $V_2$, as usual. This correspondence
sends compositions of linear mappings on \( V_1 \) to the analogous compositions of complex-linear mappings on \( V_2 \), as in Section 15.10. Thus \( \mathcal{L}(V_1) \) corresponds to a subalgebra of \( \mathcal{L}_C(V_2) \), as an algebra over \( \mathbb{R} \). It follows that \( \mathcal{L}_C(V_2) \) can be identified with the complexification of \( \mathcal{L}(V_1) \) as an algebra over \( \mathbb{R} \) in the same way, as in the preceding paragraph.

Similarly, \( gl(V_1) \) is a Lie algebra over the real numbers, and the space \( gl(V_2) = gl_C(V_2) \) of complex-linear mappings from \( V_2 \) into itself is a Lie algebra over \( \mathbb{C} \), with respect to the usual commutator brackets. As before, \( gl_C(V_2) \) can be identified with the complexification of \( gl(V_1) \), as a vector space over \( \mathbb{R} \), and in fact as a Lie algebra over \( \mathbb{R} \).

Suppose that \( V_1 \) has finite dimension as a vector space over \( \mathbb{R} \), so that \( V_2 \) has the same dimension as a vector space over \( \mathbb{C} \). Let \( \phi_1 \) be a linear mapping from \( V_1 \) into itself, and let \( \phi_2 \) be the extension of \( \phi_1 \) to a complex-linear mapping from \( V_2 \) into itself. It is easy to see that

\[
\text{tr}_{V_1} \phi_1 = \text{tr}_{V_2} \phi_2,
\]

where the left side is the trace of \( \phi_1 \) on \( V_1 \), and the right side is the trace of \( \phi_2 \) on \( V_2 \). This uses the fact that a basis for \( V_1 \) as a vector space over \( \mathbb{R} \) may be considered as a basis for \( V_2 \) as a vector space over \( \mathbb{C} \). It follows that the space \( sl(V_2) = sl_C(V_2) \) of complex-linear mappings from \( V_2 \) into itself with trace 0 corresponds to the complexification of \( sl(V_1) \) as a vector space over \( \mathbb{R} \), and hence as a Lie algebra over \( \mathbb{R} \).

Let \( n \) be a positive integer, and remember that the spaces \( M_n(\mathbb{R}), M_n(\mathbb{C}) \) of \( n \times n \) matrices with entries in \( \mathbb{R}, \mathbb{C} \), respectively, are associative algebras over \( \mathbb{R}, \mathbb{C} \) with respect to matrix multiplication. As before, \( M_n(\mathbb{C}) \) can be identified with the complexification of \( M_n(\mathbb{R}) \) as a vector space over \( \mathbb{R} \), and more precisely as an algebra over \( \mathbb{R} \). Similarly, \( gl_n(\mathbb{C}) \) can be identified with the complexification of \( gl_n(\mathbb{R}) \) as a Lie algebra over \( \mathbb{R} \). We can also identify \( sl_n(\mathbb{C}) \) with the complexification of \( sl_n(\mathbb{R}) \) as a Lie algebra over \( \mathbb{R} \).

### 15.13 Conjugate-linear involutions

Let \( A \) be a vector space over the complex numbers, and let

\[
a \mapsto a^* \tag{15.13.1}
\]

be a conjugate-linear mapping from \( A \) into itself. Put

\[
A_{sa} = \{ a \in A : a^* = a \} \tag{15.13.2}
\]

and

\[
A_{asa} = \{ a \in A : a^* = -a \} \tag{15.13.3}
\]

which are the collections of vectors in \( A \) that are \textit{self-adjoint} and \textit{anti-self-adjoint} with respect to (15.13.1), respectively. Clearly \( A_{sa} \) and \( A_{asa} \) are real-linear subspaces of \( A \), with

\[
A_{asa} = i A_{sa} \tag{15.13.4}
\]
We also have that
\[(15.13.5) \quad A_{sa} \cap A_{asa} = \{0\},\]
so that \(A_{sa}\) and \(A_{asa}\) are totally real in \(A\). Thus \(A_{sa} + A_{asa}\) can be identified with the complexifications of \(A_{sa}\) and \(A_{asa}\), as vector spaces over the real numbers, as in Section 15.10.

Suppose that
\[(15.13.6) \quad (a^*)^* = a\]
for every \(a \in A\), so that \(a \mapsto a^*\) is a conjugate-linear involution on \(A\). If \(a \in A\), then
\[(15.13.7) \quad a_{sa} = (a + a^*)/2 \in A_{sa}, \quad a_{asa} = (a - a^*)/2 \in A_{asa},\]
and \(a = a_{sa} + a_{asa}\). This means that \(A = A_{sa} + A_{asa}\), so that \(A\) can be identified with the complexifications of \(A_{sa}\) and \(A_{asa}\), as vector spaces over the real numbers.

Suppose now that \(A\) is an associative algebra over the complex numbers, and that (15.13.1) is a conjugate-linear algebra involution on \(A\). Let \(A_{\text{Lie},\mathbb{C}}\) be \(A\) considered as a Lie algebra over the complex numbers with respect to the corresponding commutator bracket, and let \(A_{\text{Lie},\mathbb{R}}\) be \(A_{\text{Lie},\mathbb{C}}\) considered as a Lie algebra over the real numbers. In this situation, \(A_{asa}\) is a Lie subalgebra of \(A_{\text{Lie},\mathbb{R}}\), and \(A_{\text{Lie},\mathbb{C}}\) can be identified with the complexification of \(A_{asa}\), as a Lie algebra over the real numbers.

Let \(V\) be a finite-dimensional vector space over the complex numbers, and let \(\beta(\cdot, \cdot)\) be a nondegenerate Hermitian form on \(V\). If \(T\) is a linear mapping from \(V\) into itself, then there is a unique adjoint linear mapping \(T^{*\beta}\) from \(V\) into itself such that
\[(15.13.8) \quad \beta(T(v), w) = \beta(v, T^{*\beta}(w))\]
for every \(v, w \in V\). Under these conditions,
\[(15.13.9) \quad T \mapsto T^{*\beta}\]
defines a conjugate-linear algebra involution on the algebra \(\mathcal{L}(V)\) of all linear mappings from \(V\) into itself. Remember that \(gl(V)\) is the same as \(\mathcal{L}(V)\), considered as a Lie algebra over the complex numbers with respect to the corresponding commutator bracket. Let \(u_\beta(V)\) be the collection of linear mappings \(T\) from \(V\) into itself that are anti-self-adjoint with respect to \(\beta\). Thus \(u_\beta(V)\) is a real-linear subspace of \(gl(V)\), and a Lie subalgebra of \(gl(V)\), considered as a Lie algebra over the real numbers. As in the preceding paragraph, \(gl(V)\) can be identified with the complexification of \(u_\beta(V)\), as a Lie algebra over the real numbers.

If \(T \in \mathcal{L}(V)\), then
\[(15.13.10) \quad \text{tr}_V T^{*\beta} = \overline{\text{tr}_V T},\]
where \(\text{tr}_V T\) denotes the trace of \(T\) on \(V\). This can be seen by choosing a basis for \(V\) to reduce to the case where \(V = \mathbb{C}^n\) for some positive integer \(n\), and expressing \(\beta(\cdot, \cdot)\) and \(T^{*\beta}\) in terms of matrices, as in Section 3.13. In particular, this implies that \(sl(V)\) is invariant under (15.13.9). Note that \(\text{tr}_V T\) is real when
T is self-adjoint with respect to $\beta$, and imaginary when $T$ is anti-self-adjoint with respect to $\beta$, by (15.13.10). Put

\[(15.13.11)\quad su_\beta(V) = sl(V) \cap u_\beta(V),\]

which is a real-linear subspace of $sl(V)$, and in fact a Lie subalgebra of $sl(V)$, considered as a Lie algebra over the complex numbers. If $T \in sl(V)$, then $T^{*,-\beta} \in sl(V)$, and hence the self-adjoint and anti-self-adjoint parts of $T$ with respect to $\beta$ are contained in $sl(V)$ too. Thus $sl(V)$ can be identified with the complexification of $su_\beta(V)$, as a Lie algebra over the real numbers.
Chapter 16

Module homomorphisms

16.1 Bilinear actions and homomorphisms

Let $k$ be a commutative ring with a multiplicative identity element, and let $V, W$ be modules over $k$. Remember that $\text{Hom}(V, W) = \text{Hom}_k(V, W)$ is the space of homomorphisms from $V$ into $W$, as modules over $k$. This is a module over $k$ too, with respect to pointwise addition and scalar multiplication of mappings. Let $A$ be another module over $k$, and let $\rho^V, \rho^W$ be bilinear actions of $A$ on $V$ and $W$, as in Section 6.1. Also let $\phi$ be a homomorphism from $V$ into $W$, as modules over $k$. Remember that $\phi$ is said to intertwine $V, W$ when

$$\phi \circ \rho^V_a = \rho^W_a \circ \phi$$

for every $a \in A$, as in Section 6.2. In this case, we may say that $\phi$ is a homomorphism from $V$ into $W$, with respect to the bilinear actions $\rho^V, \rho^W$. Let

$$\text{Hom}^A_k(V, W) = \text{Hom}^A_k(V, V)$$

be the space of these homomorphisms from $V$ into $W$ with respect to $\rho^V, \rho^W$. It is easy to see that this is a submodule of $\text{Hom}_k(V, W)$, as a module over $k$.

Let $Z$ be another module over $k$. If $\phi$ is a homomorphism from $V$ into $W$, and $\psi$ is a homomorphism from $W$ into $Z$, as modules over $k$, then $\psi \circ \phi$ is a homomorphism from $V$ into $Z$, as modules over $k$. Suppose that $\rho^Z$ is a bilinear action of $A$ on $Z$. If $\phi$ is a homomorphism from $V$ into $W$ with respect to $\rho^V, \rho^W$, and if $\psi$ is a homomorphism from $W$ into $Z$ with respect to $\rho^W, \rho^Z$, then $\psi \circ \phi$ is a homomorphism from $V$ into $Z$ with respect to $\rho^V, \rho^Z$. Indeed, if $a \in A$, then

$$\psi \circ (\rho^V_a \circ \phi) = (\psi \circ \rho^W_a) \circ \phi = \rho^Z_a \circ (\psi \circ \phi).$$

Let us now take $V = W$, and remember that $\text{Hom}_k(V, V)$ is an associative algebra over $k$, with respect to composition of mappings. More precisely, let us use the same bilinear action $\rho^V$ of $A$ on $V$ on both the domain and range of these homomorphisms. Observe that $\text{Hom}_k^A(V, V)$ is a subalgebra of $\text{Hom}_k(V, V)$, as
in the preceding paragraph. Of course, the identity mapping $I = I_V$ on $V$ is a homomorphism from $V$ into itself, with respect to $\rho_V$. Similarly, if $t \in k$, then $tI_V$ is an element of $\text{Hom}_{k}^{0}(V,V)$.

Suppose for the moment that $k$ is a field, so that $V$ is a vector space over $k$, and the algebra of linear mappings from $V$ into itself may be denoted $\mathcal{L}(V)$. Remember that $\mathcal{L}(V)$ may be used to denote the subalgebra of $\mathcal{L}(V)$ consisting of linear mappings from $V$ into itself that are homomorphisms from $V$ into itself with respect to $\rho$, as in Section 6.14. If $V$ is irreducible with respect to $\rho$, then Schur’s lemma says that every nonzero element of $\mathcal{L}(V)$ is invertible in $\mathcal{L}(V)$. If $k$ is also algebraically closed, and $V$ has finite dimension as a vector space over $k$, then $\mathcal{L}(V)$ consists exactly of the multiples of $I_V$ by elements of $k$. This is related to Exercises 4 and 5 on p54-5 of [24], for irreducible modules over a Lie algebra over $k$.

Now let $(A,[\cdot,\cdot]_A)$ be a Lie algebra over $k$, and let us consider $A$ as a module over itself, with respect to the adjoint representation. If $x,y \in A$, then put $\text{ad}_x(y) = [x,y]_A$, as usual. Let $\phi$ be a homomorphism from $A$ into itself, as a module over $k$. Thus $\phi$ is a homomorphism from $A$ into itself, as a module over itself with respect to the adjoint representation, if and only if

\begin{equation}
(16.1.4)
\phi(\text{ad}_x(y)) = \text{ad}_x(\phi(y))
\end{equation}

for every $x,y \in A$. This is the same as saying that

\begin{equation}
(16.1.5)
\phi([x,y]_A) = [x,\phi(y)]_A
\end{equation}

for every $x,y \in A$. If $k$ is a field, then we may use $\mathcal{L}^{\text{ad}}(A)$ to denote the algebra of linear mappings from $A$ into itself, as a vector space over $k$, that are homomorphisms from $A$ into itself, as a module over itself with respect to the adjoint representation.

### 16.2 Homomorphisms and semisimplicity

Let $k$ be a field, and let $(A,[\cdot,\cdot]_A)$ be a Lie algebra over $k$. Also let $A_1,\ldots,A_n$ be ideals in $A$, and suppose that $A$ corresponds to the direct sum of the $A_j$’s. This means that every $z \in A$ can be expressed in a unique way as

\begin{equation}
(16.2.1)
z = \sum_{j=1}^{n} \pi_j(z),
\end{equation}

where $\pi_j(z) \in A_j$ for each $j = 1,\ldots,n$. In particular, this implies that $A_j \cap A_l = \{0\}$ when $j \neq l$, so that $[A_j, A_l] = \{0\}$ when $j \neq l$, as in Section 10.15. It follows that $\pi_j$ is a Lie algebra homomorphism from $A$ into $A_j$ for each $j = 1,\ldots,n$. The restriction of $\pi_j$ to $A_j$ is the same as the identity mapping on $A_j$ for each $j = 1,\ldots,n$, so that $\pi_j$ maps $A$ onto $A_j$. Observe that

\begin{equation}
(16.2.2)
\pi_j([x,y]_A) = [x,\pi_j(y)]_A
\end{equation}
for every $x, y \in A$ and $j = 1, \ldots, n$. Thus $\pi_j$ may be considered as a homomorphism from $A$ into itself, as a module over itself with respect to the adjoint representation, for each $j = 1, \ldots, n$.

Let $\phi$ be a linear mapping from $A$ into itself, and suppose that $\phi$ is a homomorphism from $A$ into itself, as a module over itself with respect to the adjoint representation. Let $x \in A_j$ and $y \in A_l$ be given, for some $j, l \in \{1, \ldots, n\}$ such that $j \neq l$. This implies that $[x, y]_A = 0$, so that

\[
\pi_j(\phi(y)) = 0.
\]

Using (16.1.5) in the first step. It follows that

\[
[x, \phi(y)]_A = \phi([x, y]_A) = 0,
\]

using (16.1.5) in the first step. It follows that

\[
\pi_j(\phi(y)) = 0.
\]

If $Z(A_j) = \{0\}$ for every $j = 1, \ldots, n$, then

\[
L_{ad}(A) \cong \bigoplus_{l=1}^n k,
\]

for every $l = 1, \ldots, n$. In this situation, one can check that the restriction of $\phi$ to $A_l$ is a homomorphism from $A_l$ into itself, as a module over itself with respect to the adjoint representation.

Suppose now that $\phi_l$ is a homomorphism from $A_l$ into itself, as a module over itself with respect to the adjoint representation, for each $l = 1, \ldots, n$. Under these conditions, one can verify that

\[
\phi = \sum_{l=1}^n \phi_l \circ \pi_l
\]

defines a homomorphism from $A$ into itself, as a module over itself with respect to the adjoint representations. If $Z(A_j) = \{0\}$ for every $j = 1, \ldots, n$, then we get that $L_{ad}(A)$ corresponds to the direct sum of $L_{ad}(A_l), 1 \leq l \leq n$. This is related to parts (a) and (b) of Exercise 5 on p55 of [24].

Suppose from now on in this section that $A_j$ is simple as a Lie algebra over $k$ for each $j = 1, \ldots, n$. In particular, this means that $Z(A_j) = \{0\}$ for every $j = 1, \ldots, n$. We also get that $A_j$ is irreducible as a module over itself with respect to the adjoint representation for every $j = 1, \ldots, n$. Suppose that $k$ is algebraically closed, and that $A_j$ has positive finite dimension as a vector space over $k$ for each $j = 1, \ldots, n$. Schur’s lemma implies that $L_{ad}(A_j)$ consists of multiples of the identity mapping on $A_j$ by elements of $k$ for each $j = 1, \ldots, n$, as in the previous section. It follows that $L_{ad}(A)$ is isomorphic to the direct sum of $n$ copies of $k$, as an associative algebra over $k$, as in the preceding paragraph. This corresponds to part (a) of Exercise 5 on p55 of [24].
16.3 Linear mappings and dimensions

Let $k$ be a field, and let $n$, $r$ be positive integers. Thus the spaces $k^n$, $k^r$ of $n$, $r$-tuples of elements of $k$ may be considered as vector spaces over $k$ with respect to coordinatewise addition and scalar multiplication, respectively. Let $T$ be a linear mapping from $k^n$ into $k^r$, and let $\lambda_l(v)$ be the $l$th component of $T(v) \in k^r$ for each $l = 1, \ldots, r$ and $v \in k^n$. Equivalently, $\lambda_l$ is a linear functional on $k^n$ for each $l = 1, \ldots, r$, and any collection of $r$ linear functionals $\lambda_1, \ldots, \lambda_r$ on $k^n$ determine a linear mapping $T$ from $k^n$ into $k^r$ in this way. Let us suppose that $T \neq 0$, to avoid trivialities, so that $\lambda_l \neq 0$ for some $l$. Let $r_0$ be the maximal number of $\lambda_l$’s that are linearly independent as linear functionals on $k^n$. We can rearrange the $\lambda_l$’s, if necessary, to get that $\lambda_1, \ldots, \lambda_{r_0}$ are linearly independent as linear functionals on $k^n$, and that $\lambda_l$ can be expressed as a linear combination of $\lambda_1, \ldots, \lambda_{r_0}$ when $l > r_0$.

Similarly, let $T_0$ the linear mapping from $k^n$ into $k^{r_0}$ that corresponds to $\lambda_1, \ldots, \lambda_{r_0}$. By construction, the kernel of $T_0$ is the same as the kernel of $T$. It is well known that

$$T_0(k^n) = k^{r_0}$$

(16.3.1)

under these conditions, and in particular that $r_0 \leq n$. The dimension of the kernel of $T_0$ is $n - r_0$, as a vector space over $k$.

Let $k_1$ be a field that contains $k$ as a subfield. Thus $k^n_1$, $k^r_1$, and $k^{r_0}_1$ may be considered as vector spaces over $k_1$, which contain $k^n$, $k^r$, and $k^{r_0}$ as subsets, respectively. Let $T_{k_1}$ be the natural extension of $T$ to a mapping from $k^n_1$ into $k^r_1$ that is linear over $k_1$, and let $\lambda_{l,k_1}$ be the natural extension of $\lambda_l$ to a linear functional on $k^n_1$ for each $l = 1, \ldots, r$. Thus $\lambda_{l,k_1}$ corresponds to the $l$th coordinate of $T_{k_1}$ for each $l = 1, \ldots, r$, as before. If $l > r_0$, then $\lambda_{l,k_1}$ can be expressed as a linear combination of $\lambda_{1,k_1}, \ldots, \lambda_{r_0,k_1}$ with coefficients in $k_1$, because of the analogous property of $\lambda_l$.

Let $T_{0,k_1}$ be the natural extension of $T_0$ to a mapping from $k^n_1$ into $k^{r_0}_1$ that is linear over $k_1$. This corresponds to $\lambda_{1,k_1}, \ldots, \lambda_{r_0,k_1}$ in the usual way. The kernel of $T_{0,k_1}$ in $k^n_1$ is the same as the kernel of $T_{k_1}$, because of the property of $\lambda_{i,k_1}$ when $l > r_0$ mentioned in the preceding paragraph. It is easy to see that

$$T_{0,k_1}(k^n_1) = k^{r_0}_1,$$

(16.3.2)

using (16.3.1). This implies that $\lambda_{1,k_1}, \ldots, \lambda_{r_0,k_1}$ are linearly independent as linear functionals on $k^n_1$, and that the dimension of the kernel of $T_{0,k_1}$ in $k^n_1$ is $n - r_0$, as a vector space over $k_1$.

16.4 Homomorphisms and dimensions

Let $k$ be a field, and let $(A, [\cdot,\cdot])$ be a Lie algebra over $k$, with positive finite dimension $n$ as a vector space over $k$. We may as well take $A = k^n$, by choosing a basis for $A$ as a vector space over $k$. As in Section 9.14, the Lie bracket on $A$
can be given by
\[(x; y)_r = \sum_{j=1}^{n} \sum_{l=1}^{n} c_{r, j; l} x_j y_l \] (16.4.1)
for every \(x, y \in k^n\), where the left side is the \(r\)th coordinate of \([x, y]\) as an element of \(k^n\), and the structure constants \(c_{r, j; l}\) are elements of \(k\) for each \(j, l, r = 1, \ldots, n\).

More precisely, the structure constants satisfy (9.14.5), (9.14.7), and (9.14.8), as before.

Let \(u_1, \ldots, u_n\) be the standard basis vectors in \(k^n\), so that the \(j\)th coordinate of \(u_l\) is equal to 1 when \(j = l\), and to 0 when \(j \neq l\). Also let \(\phi\) be a linear mapping from \(k^n\) into itself, as a vector space over \(k\). It is easy to see that \(\phi\) is a homomorphism from \(A\) into itself, as a module over itself with respect to the adjoint representation, if and only if
\[\phi([u_j, u_l]) = [u_j, \phi(u_l)]\] (16.4.2)
for every \(j, l = 1, \ldots, n\). Of course, the space \(L(A)\) of linear mappings from \(A\) into itself, as a vector space over \(k\), corresponds to the space \(M_n(k)\) of \(n \times n\) matrices with entries in \(k\) in the usual way. The space \(L^\text{ad}(A)\) of homomorphisms from \(A\) into itself, as a module over itself with respect to the adjoint representation, corresponds to a linear subspace of \(M_n(k)\) that can be described in terms of linear equations for the matrix entries, using (16.4.2).

Let \(k_1\) be a field that contains \(k\) as a subfield, so that \(k^n \subseteq k_1^n\). Let us take \(A_{k_1}\) to be \(k_1^n\), as a vector space over \(k_1\) with respect to coordinatewise addition and scalar multiplication. If \(x, y \in A_{k_1}\), then \([x, y]\) can be defined as an element of \(A_{k_1}\) as in (16.4.1), where the right side is now an element of \(k_1\) for each \(r = 1, \ldots, n\). This makes \(A_{k_1}\) a Lie algebra over \(k_1\), as before.

Of course, \(u_1, \ldots, u_n\) may be considered as the standard basis vectors in \(k_1^n\) as well. Let \(\phi\) be a linear mapping from \(k_1^n\) into itself, as a vector space over \(k_1\). As before, \(\phi\) is a homomorphism from \(A_{k_1}\) into itself, as a module over itself with respect to the adjoint representation, if and only if (16.4.2) holds for every \(j, l = 1, \ldots, n\). As usual, the space \(L(A_{k_1})\) of linear mappings from \(A_{k_1}\) into itself, as a vector space over \(K\), corresponds to the space \(M_n(k_1)\) of \(n \times n\) matrices with entries in \(k_1\). The space \(L^\text{ad}(A_{k_1})\) of homomorphisms from \(A_{k_1}\) into itself, as a module over itself with respect to the adjoint representation, corresponds to a linear subspace of \(M_n(k_1)\) that can be described in terms of essentially the same linear equations for the matrix entries as for \(k\), using (16.4.2) again.

This brings us to the same type of situation as discussed in the previous section. It follows that the dimension of \(L^\text{ad}(A_{k_1})\), as a vector space over \(k_1\), is the same as the dimension of \(L^\text{ad}(A)\), as a vector space over \(k\). This is related to part of part (b) of Exercise 5 on p55 of [24].

### 16.5 Semisimplicity and dimensions

Let \(k\) be a field of characteristic 0, and let \((A, [, ;])\) be a Lie algebra over \(k\) of positive finite dimension \(n\). As in the previous section, we may as well take
$A = k^n$, with Lie bracket as in (16.4.1). Let $k_1$ be an algebraically closed field that contains $k$ as a subfield. As before, $A_{k_1} = k_1^n$ is a Lie algebra over $k_1$, with Lie bracket as in (16.4.1). In this section, we suppose that $A$ is semisimple as a Lie algebra over $k$, which is equivalent to asking that $A_{k_1}$ be semisimple as a Lie algebra over $k_1$, as in Section 11.5.

As in Section 10.15, $A_{k_1}$ is isomorphic to the direct sum of $h$ simple Lie algebras over $k_1$, for some positive integer $h$. Remember that the space $L^{\text{ad}}(A_{k_1})$ of module homomorphisms from $A_{k_1}$ into itself, as a module over itself with respect to the adjoint representation, is an associative algebra over $k_1$ with respect to composition of mappings. In this situation, $L^{\text{ad}}(A_{k_1})$ is isomorphic to the direct sum of $h$ copies of $k_1$, as in Section 16.2. In particular, the dimension of $L^{\text{ad}}(A_{k_1})$ is $h$, as a vector space over $k_1$. We also get that $L^{\text{ad}}(A_{k_1})$ is commutative as an algebra over $k_1$.

Similarly, the space $L^{\text{ad}}(A)$ of module homomorphisms from $A$ into itself, as a module over itself with respect to the adjoint representation, is an associative algebra over $k$ with respect to composition of mappings. As in the previous section, the dimension of $L^{\text{ad}}(A)$, as a vector space over $k$, is the same as the dimension of $L^{\text{ad}}(A_{k_1})$ as a vector space over $k_1$, which is equal to $h$. This corresponds to the first part of part (b) of Exercise 5 on p55 of [24].

If $\phi$ is a linear mapping from $A$ into itself, as a vector space over $k$, then $\phi$ has a natural extension to a linear mapping from $A_{k_1}$ into itself, as a vector space over $k_1$. If $\phi$ is a homomorphism from $A$ into itself, as a module over itself with respect to the adjoint representation, then the extension of $\phi$ to $A_{k_1}$ is a homomorphism from $A_{k_1}$ into itself, as a module over itself with respect to the adjoint representation. This follows from the characterization of module homomorphisms as linear mappings that satisfy (16.4.2) in both cases. Using this, we get that $L^{\text{ad}}(A)$ is commutative as an algebra over $k$ with respect to composition of mappings, because $L^{\text{ad}}(A_{k_1})$ is commutative, as before. This is related to the second part of part (b) of Exercise 5 on p55 of [24].

Suppose that $A$ is simple as a Lie algebra over $k$, which implies that $A$ is irreducible as a module over itself, with respect to the adjoint representation. In this case, nonzero elements of $L^{\text{ad}}(A)$ are invertible in $L^{\text{ad}}(A)$, by Schur’s lemma, as in Section 16.1. This means that $L^{\text{ad}}(A)$ is a field, because $L^{\text{ad}}(A)$ is commutative, as in the preceding paragraph. This corresponds to the second part of part (b) of Exercise 5 on p55 of [24], with $m = 1$. Otherwise, $A$ is isomorphic to the direct sum of simple Lie algebras $A_1, \ldots, A_m$ over $k$ for some positive integer $m$, as in Section 10.15. This implies that $L^{\text{ad}}(A)$ is isomorphic to the direct sum of $L^{\text{ad}}(A_1), \ldots, L^{\text{ad}}(A_m)$ as an associative algebra over $k$, as in Section 16.2. It follows that $L^{\text{ad}}(A)$ is isomorphic to the direct sum of $m$ fields, as in the second part of part (b) of Exercise 5 on p55 of [24].

16.6 Absolutely simple Lie algebras

Let $k$ be a field, and let $(A, [\cdot, \cdot])$ be a Lie algebra over $k$ of positive finite dimension $n$, as a vector space over $k$. As before, we may as well take $A$ to be
$k^n$, with Lie bracket as in (16.4.1). Let $k_1$ be an algebraically closed field that contains $k$ as a subfield, and take $A_{k_1} = k_1^n$ as a Lie algebra over $k_1$, with Lie bracket defined as in (16.4.1). If $A_{k_1}$ is simple as a Lie algebra over $k_1$, then $A$ is said to be absolutely simple as a Lie algebra over $k$, as in part (c) of Exercise 5 on p55 of [24]. This implies that $A$ is simple as a Lie algebra over $k$, as in Section 11.4.

Suppose that $A$ is absolutely simple, so that $A_{k_1}$ is simple as a Lie algebra over $k_1$. This implies that $A_{k_1}$ is irreducible as a module over itself with respect to the adjoint representation. Remember that $\mathcal{L}^{ad}(A_{k_1})$ is the algebra of homomorphisms from $A_{k_1}$ into itself, as a module over itself with respect to the adjoint representation. Under these conditions, $\mathcal{L}^{ad}(A_{k_1})$ consists exactly of multiples of the identity mapping on $A_{k_1}$ by elements of $k_1$, by Schur’s lemma, as in Section 16.1. Equivalently, this means that the dimension of $\mathcal{L}^{ad}(A_{k_1})$, as a vector space over $k_1$, is equal to one. It follows that the space $\mathcal{L}^{ad}(A)$ of homomorphisms from $A$ into itself, as a module over itself with respect to the adjoint representation, is equal to one as a vector space over $k$, as in Section 16.4. Of course, $\mathcal{L}^{ad}(A)$ automatically contains all multiples of the identity mapping on $A$ by elements of $k$. This means that $\mathcal{L}^{ad}(A)$ consists exactly of multiples of the identity mapping on $A$ by elements of $k$ in this situation. This corresponds to part of the first part of part (c) of Exercise 5 on p55 of [24].

Conversely, suppose that $\mathcal{L}^{ad}(A)$ consists exactly of multiples of the identity mapping on $A$ by elements of $k$. This is the same as saying that $\mathcal{L}^{ad}(A)$ has dimension one as a vector space over $k$, as before. It follows that $\mathcal{L}^{ad}(A_{k_1})$ has dimension one as a vector space over $k_1$, as in Section 16.4. Suppose that $k$ has characteristic 0, so that $k_1$ has characteristic 0 too. If $A_{k_1}$ is semisimple as a Lie algebra over $k_1$, then $A_{k_1}$ is isomorphic to the direct sum of $h$ simple Lie algebras over $k_1$, for some positive integer $h$, as in Section 10.15. This implies that the dimension of $\mathcal{L}^{ad}(A_{k_1})$ is equal to $h$ as a vector space over $k_1$, as in the previous section. Under these conditions, we get that $h = 1$, so that $A_{k_1}$ is simple as a Lie algebra over $k_1$. This corresponds to the other part of the first part of part (c) of Exercise 5 on p55 of [24].

Suppose now that $A$ is simple as a Lie algebra over $k$, and that $k$ has characteristic 0. In particular, this means that $A$ is semisimple as a Lie algebra over $k$. Put $K = \mathcal{L}^{ad}(A)$, which is a field in this situation, as in the previous section. By construction, the elements of $K$ are linear mappings from $A$ into itself, as a vector space over $k$. We may consider $A$ as a vector space over $K$, where scalar multiplication by an element of $K$ is defined by the corresponding linear mapping on $A$. It is easy to see that the Lie bracket on $A$ is bilinear over $K$, as a mapping from $A \times A$ into $A$, because the elements of $K$ are homomorphisms from $A$ into itself, as a module over itself as a Lie algebra over $k$, with respect to the adjoint representation. Thus $A$ may be considered as a Lie algebra over $K$, with respect to the Lie bracket already defined on $A$ as a Lie algebra over $k$.

Let us use $A_K$ to refer to $A$ as a Lie algebra over $K$ in this way. As before, $\mathcal{L}^{ad}(A_K)$ denotes the space of linear mappings from $A_K$ into itself, as a vector space over $K$, that are homomorphisms from $A_K$ into itself as a module over itself, as a Lie algebra over $K$, and with respect to the adjoint representation.
Let $k$ be a commutative ring with a multiplicative identity element, and let $k_0$ be a subring of $k$ that contains the multiplicative identity element. If $A$ is a module over $k$, then $A$ may be considered as a module over $k_0$ as well. Let $A_0$ be $A$ considered as a module over $k_0$. If $B$ is a submodule of $A$, as a module over $k$, then $B$ may be considered as a submodule of $A_0$ too.

Similarly, if $A$ is an algebra over $k$ in the strict sense, then $A$ may be considered as an algebra over $k_0$ in the strict sense. Let $A_0$ be $A$ considered as an algebra over $k_0$ in the strict sense. If $B$ is a subalgebra of $A$, as an algebra over $k$ in the strict sense, then $B$ may be considered as a subalgebra of $A_0$, as an algebra over $k_0$ in the strict sense. If $B$ is a one or two-sided ideal in $A$, as an algebra over $k$ in the strict sense, then $B$ has the analogous property in $A_0$ as well.

Now let $(A, \{\cdot, \cdot\})$ be a Lie algebra over $k$. As before, let $A_0$ be $A$, considered as a Lie algebra over $k_0$. It is easy to see that $A$ is commutative, solvable, or nilpotent as a Lie algebra over $k$ if and only if $A_0$ has the same property as a Lie algebra over $k_0$.

Let $B$ be an ideal in $A$, as a Lie algebra over $k$. Thus $B$ may be considered as an ideal in $A_0$, as a Lie algebra over $k_0$. Let $B_0$ be $B$, considered as an ideal in $A_0$, and as a Lie algebra over $k_0$ in particular. If $B$ is solvable, as a Lie algebra over $k$, then $B_0$ is solvable, as a Lie algebra over $k_0$, as in the preceding paragraph. If $A_0$ is semisimple as a Lie algebra over $k_0$, then it follows that $A$ is semisimple as a Lie algebra over $k$.

Of course, $k$ may be considered as a module over $k_0$. Let us suppose from now on in this section that $k$ is free as a module over $k_0$, of rank $n_0$ for some positive integer $n_0$. This means that $k$ is isomorphic to $k_0^{n_0}$ as a module over $k_0$, where $k_0^{n_0}$ is the space of $n_0$-tuples of elements of $k_0$, considered as a module over $k_0$ with respect to pointwise addition and scalar multiplication.

If $x \in k$, then $M_x(y) = x y$ defines a homomorphism from $k$ into itself, as a module over $k_0$. Let $tr_{k/k_0} x$ be the trace of $M_x$, as a homomorphism from $k$ into itself, as a free module over $k_0$ of rank $n_0$. This defines a homomorphism
CHAPTER 16. MODULE HOMOMORPHISMS

from $k$ into $k_0$, as modules over $k_0$. This uses the fact that $x \mapsto M_x$ is linear over $k_0$, as a mapping from $k$ into the space homomorphisms from $k$ into itself, as a module over $k_0$. If $x \in k_0$, then

$$\text{tr}_{k/k_0} x = n_0 \cdot x.$$  

(16.7.1)

Let $n$ be a positive integer, so that $V = k^n$ is a free module over $k$ with respect to coordinatewise addition and scalar multiplication. We may also consider $V$ as a free module over $k_0$ of rank $n_0 n$, because $k$ is a free module over $k_0$ of rank $n_0$. If $T$ is a homomorphism from $V$ into itself as a module over $k$, then $T$ may be considered as a homomorphism from $V$ into itself as a module over $k_0$ as well. Let $\text{tr}_{V,k} T \in k$ be the trace of $T$ as a homomorphism from $V$ into itself as a free module over $k$ of rank $n_0$, and let $\text{tr}_{V,k_0} T \in k_0$ be the trace of $T$ as a homomorphism from $V$ into itself as a free module over $k_0$ of rank $n_0 n$.

One can check that

$$\text{tr}_{V,k_0} T = \text{tr}_{k/k_0} (\text{tr}_{V,k} T).$$

(16.7.2)

16.8 Lie algebras over subfields

Let $k$ be a field, and let $k_0$ be a subfield of $k$. Also let $(A, [, [, ]])$ be a Lie algebra over $k$, and let $A_0$ be $A$ considered as a Lie algebra over $k_0$. Suppose that $k$ has finite dimension $n_0$ as a vector space over $k_0$, and that $A$ has positive finite dimension $n$ as a vector space over $k$. Thus $A_0$ has dimension $n_0 n$ as a vector space over $k_0$. If $x \in A$, then put $\text{ad}_x(z) = [x, z]$ for every $z \in A$, as usual. This defines $\text{ad}_x$ as a linear mapping from $A$ into itself, as a vector space over $k$. Of course, $\text{ad}_x$ may be considered as a linear mapping from $A_0$ into itself as well, as a vector space over $k_0$.

The Killing form for $A$, as a Lie algebra over $k$, is defined by

$$b(x, y) = \text{tr}_A(\text{ad}_x \circ \text{ad}_y) \in k$$

(16.8.1)

for every $x, y \in A$, using the trace of $\text{ad}_x \circ \text{ad}_y$ on $A$ as a vector space over $k$. Similarly, the Killing form for $A_0$ is defined by

$$b_0(x, y) = \text{tr}_{A_0}(\text{ad}_x \circ \text{ad}_y) \in k_0$$

(16.8.2)

for every $x, y \in A_0$, using the trace of $\text{ad}_x \circ \text{ad}_y$ on $A_0$ as a vector space over $k$. Let $\text{tr}_{k/k_0}$ be the trace mapping from $k$ into $k_0$ mentioned in the previous section. If $x, y \in A$, then

$$b_0(x, y) = \text{tr}_{k/k_0} b(x, y),$$

(16.8.3)

as in (16.7.2).

Suppose that $b(\cdot, \cdot)$ is nondegenerate on $A$. Let $x \in A$ with $x \neq 0$ be given, so that there is a $y \in A$ such that $b(x, y) \neq 0$. Put $y_1 = (1/b(x, y)) y$, so that

$$b(x, y_1) = 1.$$  

(16.8.4)
This implies that
\[ b_0(x, y_1) = \text{tr}_{k/k_0} 1 = n_0 \cdot 1, \tag{16.8.5} \]
using (16.8.3) in the first step, and (16.7.1) in the second step.

Suppose from now on in this section that \( k_0 \) has characteristic 0, which is the same as saying that \( k \) has characteristic 0. In this case, (16.8.5) implies that \( b_0(x, y_1) \neq 0 \) in \( k_0 \). This means that \( b_0(\cdot, \cdot) \) is nondegenerate on \( A_0 \) in this situation.

Suppose that \( A \) is semisimple as a Lie algebra over \( k \), so that \( b(\cdot, \cdot) \) is non-degenerate on \( A \), as in Section 10.13. This implies that \( b_0(\cdot, \cdot) \) is nondegenerate on \( A_0 \), as in the preceding paragraphs. It follows that \( A_0 \) is semisimple as a Lie algebra over \( k_0 \), as in Section 10.13 again.

As in Section 10.15, there are finitely many ideals \( A_{0,1}, \ldots, A_{0,r} \) in \( A_0 \), as a Lie algebra over \( k_0 \), such that \( A_{0,j} \) is simple as a Lie algebra over \( k_0 \) for each \( j = 1, \ldots, r \), and \( A_0 \) corresponds to the direct sum of \( A_{0,1}, \ldots, A_{0,r} \) as a Lie algebra over \( k_0 \). If \( t \in k \), then
\[ \phi_t(x) = tx \tag{16.8.6} \]
defines a linear mapping from \( A_0 \) into itself as a vector space over \( k_0 \), and in fact a homomorphism from \( A_0 \) into itself as a module over itself, with respect to the adjoint representation. It follows that
\[ \phi_t(A_{0,j}) \subseteq A_{0,j} \tag{16.8.7} \]
for each \( j = 1, \ldots, r \), as in Section 16.2. This means that \( A_{0,j} \) is a linear subspace of \( A \), as a vector space over \( k \), for each \( j = 1, \ldots, r \). Thus \( A_{0,j} \) is an ideal in \( A \), as a Lie algebra over \( k \), for each \( j = 1, \ldots, r \).

In particular, \( A_{0,j} \) is a Lie subalgebra of \( A \), as a Lie algebra over \( k \), for each \( j = 1, \ldots, r \). It is easy to see that \( A_{0,j} \) is simple as a Lie algebra over \( k \) for each \( j = 1, \ldots, r \), because \( A_{0,j} \) is simple as a Lie algebra over \( k_0 \).

If \( A \) is simple as a Lie algebra over \( k \), then \( A \) is semisimple. In this case, we get that \( r = 1 \) in the previous argument, so that \( A_0 = A_{0,1} \) is simple as a Lie algebra over \( k_0 \). This corresponds to part (d) of Exercise 5 on p55 of [24].
Part III

Toral subalgebras and roots
Chapter 17

Subalgebras and diagonalizability

17.1 Toral subalgebras

Let $k$ be a field, and let $(A, [\cdot, \cdot]_A)$ be a finite-dimensional Lie algebra over $k$. Suppose that $y, z$ are ad-diagonalizable elements of $A$ such that

\[(17.1.1) \quad [y, z]_A = 0.\]

This implies that $\text{ad}_y$ and $\text{ad}_z$ commute as linear mappings from $A$ into itself, as in Section 2.4. It follows that $\text{ad}_y + \text{ad}_z$ is diagonalizable as a linear mapping from $A$ into itself, as in Section 10.6. This means that $y + z$ is ad-diagonalizable, as an element of $A$.

If every element of $A$ is ad-nilpotent, then $A$ is nilpotent as a Lie algebra over $k$, as in Section 9.10. Suppose that $k$ is an algebraically closed field of characteristic 0, and that $A$ is semisimple as a Lie algebra over $k$. If $A \neq \{0\}$, then $A$ is not nilpotent, and hence there is an $x \in A$ that is not ad-nilpotent. Using the abstract Jordan decomposition, as in Section 14.3, we get that there are $x_1, x_2 \in A$ such that $x = x_1 + x_2$, $x_1$ is ad-diagonalizable, and $x_2$ is ad-nilpotent. In particular, $x_1 \neq 0$, because $x$ is not ad-nilpotent.

A Lie subalgebra $B$ of $A$ is said to be toral if every element of $B$ is ad-diagonalizable, as an element of $A$, at least when $k$ is algebraically closed, as on p35 of [13]. Under the conditions mentioned in the preceding paragraph, there are nonzero toral subalgebras of $A$, as on p35 of [13]. More precisely, one can take the linear span of a nonzero ad-diagonalizable element of $A$.

Let $k$ be any field again, and let $A$ be a finite-dimensional Lie algebra over $k$. If $B$ is a Lie subalgebra of $A$, and if every element of $B$ is ad-diagonalizable as an element of $A$, then $B$ is commutative as a Lie algebra over $k$, as in the lemma on p35 of [13]. To see this, let $x \in B$ be given, and note that $\text{ad}_x$ maps $B$ into itself, because $B$ is a Lie subalgebra of $A$. The restriction of $\text{ad}_x$ to $B$ is the same as $\text{ad}_{B,x}$. We would like to show that $\text{ad}_{B,x} = 0$.  

335
CHAPTER 17. SUBALGEBRAS AND DIAGONALIZABILITY

Of course, $\text{ad}_x$ is diagonalizable as a linear mapping from $A$ into itself, by hypothesis. This implies that $\text{ad}_{B,x}$ is diagonalizable as a linear mapping from $B$ into itself, as in Section 10.6. Let $\lambda \in k$ be an eigenvalue of $\text{ad}_{B,x}$, so that there is a $y \in B$ such that $y \neq 0$ and

\begin{equation}
\text{ad}_{B,x}(y) = [x,y]_A = \lambda y.
\end{equation}

We would like to show that $\lambda = 0$.

Equivalently,

\begin{equation}
\text{ad}_{B,y}(x) = -\lambda y,
\end{equation}

by (17.1.2). Remember that $y$ is ad-diagonalizable as an element of $A$, by hypothesis. This implies that $\text{ad}_{B,y}$ is diagonalizable as a linear mapping from $B$ into itself, as before. Thus $x$ can be expressed as a sum of eigenvectors of $\text{ad}_{B,y}$. It follows that $\text{ad}_{B,y}(x)$ is a sum of eigenvectors of $\text{ad}_{B,y}$ corresponding to nonzero eigenvalues of $\text{ad}_{B,y}$, if there are any.

If $\text{ad}_{B,y}(x) \neq 0$, then the previous statement implies that $\text{ad}_{B,y}(\text{ad}_{B,y}(x)) \neq 0$. However,

\begin{equation}
\text{ad}_{B,y}(\text{ad}_{B,y}(x)) = -\lambda \text{ad}_{B,y}(y) = -\lambda [y,y]_A = 0,
\end{equation}

using (17.1.3) in the first step. Thus $\text{ad}_{B,y}(x) = 0$, so that $\lambda = 0$, by (17.1.3). This means that $\text{ad}_{B,x} = 0$, because $\text{ad}_{B,x}$ is diagonalizable on $B$, and 0 is its only eigenvalue. This shows that $B$ is commutative as a Lie algebra, as desired.

Note that we only used the diagonalizability of $\text{ad}_{B,z}$ on $B$ for each $z \in B$, rather than the diagonalizability of $\text{ad}_z$ on $A$. This also corresponds to simply taking $A = B$.

17.2 Simultaneous diagonalizability

Let $k$ be a field, and let $(A, [\cdot, \cdot]_A)$ be a finite-dimensional Lie algebra over $k$. Also let $B$ be a Lie subalgebra of $A$, and suppose that every element of $B$ is ad-diagonalizable as an element of $A$. Thus $B$ is commutative as a Lie algebra over $k$, as in the previous section. Let $B'$ be the dual of $B$, as a vector space over $k$. Remember that $B'$ has the same dimension as $B$.

Suppose that $x \in A$ is an eigenvector for $\text{ad}_w$ for each $w \in B$, so that

\begin{equation}
\text{ad}_w(x) = [w,x]_A = \alpha(w) x
\end{equation}

for some $\alpha(w) \in k$. If $x \neq 0$, then $\alpha(w)$ is linear in $w$, and hence defines an element of $B'$.

If $\alpha \in B'$, then put

\begin{equation}
A_\alpha = \{ x \in A : \text{ for each } w \in B, \text{ ad}_w(x) = [w,x]_A = \alpha(w) x \}.
\end{equation}

This is a linear subspace of $A$, as a vector space over $k$. An element of $A_\alpha$ is said to have weight $\alpha$, as on p43 of [23].
In particular, if we take \( \alpha = 0 \), then we get
\[
A_0 = \{ x \in A : \text{ for each } w \in B, \ \text{ad}_w(x) = [w, x]_A = 0 \}.
\]
This is the same as the centralizer \( C_A(B) \) of \( B \) in \( A \), as in Section 7.6. Note that
\[
B \subseteq A_0 = C_A(B),
\]
because \( B \) is commutative as a Lie algebra, as before.

Let us suppose that \( B \neq \{0\} \), to avoid trivialities. This implies that \( A_0 \neq \{0\} \), by (17.2.4). Put
\[
\Phi_B = \{ \alpha \in B' : \alpha \neq 0 \text{ and } A_\alpha \neq \{0\} \}.
\]
Thus \( \Phi_B \cup \{0\} \) is the same as the set of \( \alpha \in B' \) such that \( A_\alpha \neq \{0\} \).

If \( u, v \in B \), then \( [u, v]_A = 0 \), as before. This implies that \( \text{ad}_u \) and \( \text{ad}_v \)
commute as linear mappings from \( A \) into itself, as in Section 2.4. By hypothesis, \( \text{ad}_u \)
is diagonalizable as a linear mapping from \( A \) into itself for each \( u \in B \). It follows that the linear mappings \( \text{ad}_u, u \in B \), are simultaneously diagonalizable
on \( A \), by standard arguments.

This means that \( A \) corresponds to the direct sum of the subspaces \( A_\alpha \) with \( \alpha \in \Phi_B \cup \{0\} \), as a vector space over \( k \). In particular, the number of elements of \( \Phi_B \) is strictly less than the dimension of \( A \). This corresponds to some of the
remarks on p35 of [13], and to Theorem 1 on p43 of [23].

Let \( \alpha, \beta \in B' \) be given, and suppose that \( x \in A_\alpha, y \in A_\beta \). If \( w \in B \), then
\[
[w, [x, y]_A]_A = [[w, x]_A, y]_A + [x, [w, y]_A]_A
= \alpha(w) [x, y]_A + \beta(w) [x, y]_A.
\]
This uses the Jacobi identity in the first step, or, equivalently, the fact that \( \text{ad}_w \)
is a derivation on \( A \). Thus
\[
[x, y]_A \in A_{\alpha+\beta}
\]
under these conditions. This corresponds to the first part of the proposition near the top of p36 in [13], and to the statement 2.1 on p45 of [23].

If \( \alpha \in B' \) and \( x \in A_\alpha \), then
\[
\text{ad}_x(A_\gamma) \subseteq A_{\alpha+\gamma}
\]
for every \( \gamma \in B' \), by (17.2.7). This implies that
\[
(\text{ad}_x)^n(A_\gamma) \subseteq A_{\alpha+\gamma}
\]
for every positive integer \( n \), where \( (\text{ad}_x)^n \) is the \( n \)th power of \( \text{ad}_x \) on \( A \), with respect to composition of mappings. If \( \alpha \neq 0 \) and \( k \) has characteristic 0, then one can use this to get that \( x \) is ad-nilpotent as an element of \( A \), as in the second
part of the proposition at the top of p36 in [13]. This could also be obtained as in Section 14.2.

Similarly, if \( x \in A_\alpha \) and \( y \in A_\beta \) for some \( \alpha, \beta \in B' \), then
\[
(\text{ad}_x \circ \text{ad}_y)(A_\gamma) = \text{ad}_x(\text{ad}_y(A_\gamma)) \subseteq A_{\alpha+\beta+\gamma}
\]
for every \( \gamma \in B' \), by (17.2.8).
17.3 Related bilinear forms

Let us continue with the same notation and hypotheses as in the previous section. Let \( b(\cdot, \cdot) \) be a bilinear form on \( A \) that is associative, or equivalently invariant with respect to the adjoint representation on \( A \). Thus

\[
b([w, x]_A, y) = -b(x, [w, y]_A)
\]

for every \( w, x, y \in A \), as in Sections 6.10 and 7.7. In particular, the Killing form

\[
b_A(x, y) = \text{tr}_A(\text{ad}_x \circ \text{ad}_y)
\]

on \( A \) has this property, as in Section 7.9.

Let \( \alpha, \beta \in B' \), \( x \in A_\alpha, \ y \in A_\beta \), and \( w \in B \) be given. Using (17.3.1), we get

\[
\alpha(w) b(x, y) = -\beta(w) b(x, y).
\]

If \( \alpha + \beta \neq 0 \), then there is a \( w \in B \) such that \( \alpha(w) \neq -\beta(w) \), and hence

\[
b(x, y) = 0.
\]

This corresponds to the third part of the proposition on the top of p36 in [13], and the first part of Theorem 3 (i) on p44 of [23]. Alternatively, if \( b(\cdot, \cdot) \) is the Killing form (17.3.2) on \( A \), then (17.3.4) follows from (17.2.10).

Let \( \alpha \in B' \) and \( w \in B \) be given again. If \( x \in A_\alpha \) and \( y \in A_\beta \), then

\[
b(w, [x, y]_A) = b([w, x]_A, y) = \alpha(w) b(x, y).
\]

Similarly, if \( x \in A \) and \( y \in A_{-\alpha} \), then

\[
b([x, y]_A, w) = b(x, [y, w]_A) = \alpha(w) b(x, y).
\]

This corresponds to Theorem 3 (iii) on p44 of [23], and is related to part (c) of the proposition on p37 of [13]. Note that

\[
[x, y]_A \in A_0
\]

when \( x \in A_\alpha \) and \( y \in A_{-\alpha} \), as in (17.2.7).

Suppose now that \( b(\cdot, \cdot) \) is also nondegenerate on \( A \). In this case, the restriction of \( b(\cdot, \cdot) \) to \( A_0 \) is nondegenerate on \( A_0 \), because of (17.3.4). This corresponds to the corollary on p36 of [13], and to the third part of Theorem 3 (i) on p44 of [23].

Let \( \alpha \in B' \) with \( \alpha \neq 0 \) be given. If \( x \in A_\alpha \) and \( x \neq 0 \), then there is a \( y \in A \) such that

\[
b(x, y) \neq 0,
\]

because \( b(\cdot, \cdot) \) is nondegenerate on \( A \). More precisely, we can take \( y \in A_{-\alpha} \), because of (17.3.4). In particular, this shows that \( A_{-\alpha} \neq \{0\} \) when \( A_\alpha \neq \{0\} \). Equivalently, if \( \alpha \in \Phi_B \), then \( -\alpha \in \Phi_B \). This corresponds to part (b) of the proposition on p37 of [13], and part of Theorem 2 (a) on p43 of [23]. Similarly, the restriction of \( b(\cdot, \cdot) \) to

\[
A_\alpha + A_{-\alpha}
\]

is nondegenerate. This corresponds to the second part of Theorem 3 (i) on p44 of [23].
17.4 Maximal toral subalgebras

Let $k$ be a field, and let $V$ be a finite-dimensional vector space over $k$. Also let $R$ and $T$ be commuting linear mappings from $V$ into itself. If $T$ is nilpotent on $V$, then $R \circ T$ is nilpotent on $V$ as well. In particular, this implies that

$$\text{tr}_V (R \circ T) = 0.$$  
(17.4.1)

This is the lemma on p36 of [13].

Let $k$ be an algebraically closed field of characteristic 0, and let $(A, [\cdot, \cdot]_A)$ be a finite-dimensional semisimple Lie algebra over $k$. Suppose that $B$ is a toral subalgebra of $A$, as in Section 17.1, and that $B$ is maximal with respect to inclusion. Under these conditions,

$$C_A(B) = B,$$  
(17.4.2)

where $C_A(B)$ is the centralizer of $B$ in $A$, as in Section 7.6. This is the second proposition on p36 of [13].

Remember that $C_A(B)$ is a Lie subalgebra of $A$, as in Section 7.6. In this situation, $B \subseteq C_A(B)$, because $B$ is commutative as a Lie algebra, as in (17.2.4). Thus we would like to show that

$$C_A(B) \subseteq B.$$  
(17.4.3)

Let $x \in C_A(B)$ be given, and let

$$x = x_1 + x_2$$  
(17.4.4)

be the abstract Jordan decomposition of $x$ in $A$, as in Section 14.3. Thus $x_1, x_2 \in A$,

$$\text{ad}_x = \text{ad}_{x_1} + \text{ad}_{x_2},$$  
(17.4.5)

and $\text{ad}_{x_1}, \text{ad}_{x_2}$ are the diagonalizable and nilpotent parts of $\text{ad}_x$, as a linear mapping from $A$ into itself, as in Section 10.8. The condition that $x \in C_A(B)$ means exactly that $\text{ad}_x$ maps $B$ into $\{0\}$. This implies that $\text{ad}_{x_1}$ and $\text{ad}_{x_2}$ map $B$ into $\{0\}$ as well, as in Section 10.8. It follows that

$$x_1, x_2 \in C_A(B),$$  
(17.4.6)

as in Step (1) of the proof on p36 of [13].

If $x \in C_A(B)$, then the linear span

$$B(x) = \{w + tx : w \in B, t \in k\}$$  
(17.4.7)

of $B$ and $x$ in $A$ is a Lie subalgebra of $A$ that is commutative as a Lie algebra. More precisely, this uses the fact that $B$ is commutative as a Lie algebra, and that $[w, x]_A = 0$ for every $w \in B$. If $x$ is ad-diagonalizable as an element of $A$ too, then $B(x)$ is a toral subalgebra of $A$, by the remark at the beginning of
Section 17.1. This implies that $B(x) = B$, because $B$ is supposed to be maximal in $A$, so that
\begin{equation}
(17.4.8)
x \in B.
\end{equation}
This is Step (2) of the proof on p36 of [13].

Let $b_A(\cdot, \cdot)$ be the Killing form on $A$, as in (17.3.2). Remember that $b_A(\cdot, \cdot)$ is nondegenerate on $A$, because $A$ is semisimple and $k$ has characteristic 0, as in Section 10.13. This implies that the restriction of $b_A(\cdot, \cdot)$ to $A_0 = C_A(B)$ is nondegenerate, as in the previous section. Step (3) of the proof on p36 of [13] states that the restriction of $b_A(\cdot, \cdot)$ to $B$ is nondegenerate.

To see this, let $x \in B$ be given, and suppose that $b_A(x, y) = 0$ (17.4.9) for every $y \in C_A(B)$. We would like to show that $x = 0$. If (17.4.9) holds for every $y \in C_A(B)$, then $x = 0$, because $b_A(\cdot, \cdot)$ is nondegenerate on $C_A(B)$, as in the preceding paragraph.

Let $z \in C_A(B)$ be given, and so that $[x, z]_A = 0$, because $x \in B$. This implies that $\text{ad}_x$ and $\text{ad}_z$ commute as linear mappings from $A$ into itself, as in Section 2.4. If $\text{ad}_z$ is nilpotent on $A$, then it follows that
\begin{equation}
(17.4.10)
b_A(x, z) = \text{tr}_A(\text{ad}_x \circ \text{ad}_z) = 0,
\end{equation}
as in (17.4.1).

Let $y \in C_A(B)$ be given, and let $y = y_1 + y_2$ be the abstract Jordan decomposition of $y$. Thus $y_1, y_2 \in C_A(B)$, as before. We also have that $y_1 \in B$, because $y_1$ is ad-diagonalizable as an element of $A$. It follows that $b_A(x, y_1) = 0$, by hypothesis. Observe that $b_A(x, y_2) = 0$, because $y_2$ is ad-nilpotent as an element of $A$, as in (17.4.10). Combining these two statements, we get that (17.4.9) holds. This implies that $x = 0$, because $y$ is an arbitrary element of $C_A(B)$, as desired.

Step (4) of the proof on p36 of [13] states that $C_A(B)$ is nilpotent as a Lie algebra over $k$. Let $x \in C_A(B)$ be given. We would like to show that $x$ is ad-nilpotent as an element of $C_A(B)$, which is to say that $\text{ad}_{C_A(B), x}$ is nilpotent as a linear mapping from $C_A(B)$ into itself. Let (17.4.4) be the abstract Jordan decomposition of $x$ in $A$ again. Thus $x_1, x_2 \in B$, by the first two steps, which implies that $\text{ad}_{C_A(B), x_1} = 0$. We also have that $\text{ad}_{C_A(B), x_2}$ is nilpotent on $C_A(B)$, because $\text{ad}_{x_2}$ is nilpotent on $A$ by construction, and $\text{ad}_{C_A(B), x_2}$ is the same as the restriction of $\text{ad}_{x_2}$ to $C_A(B)$. This implies that
\begin{equation}
(17.4.11)
\text{ad}_{C_A(B), x} = \text{ad}_{C_A(B), x_2}
\end{equation}
is nilpotent on $C_A(B)$. It follows that $C_A(B)$ is nilpotent as a Lie algebra over $k$, as in Section 9.10.

Step (5) of the continuation of the proof on p37 of [13] states that
\begin{equation}
(17.4.12)
B \cap ([C_A(B), C_A(B)]) = \{0\}.
\end{equation}
To see this, observe that

\[(17.4.13) \quad b_A([x_1, x_2]_A, y) = b_A(x_1, [x_2, y]_A) = 0\]

for every \(x_1, x_2 \in C_A(B)\) and \(y \in B\). This uses the associativity of \(b_A(\cdot, \cdot)\) on \(A\) in the first step, and the fact that \([x_2, y]_A = 0\) in the second step. It follows that \(b_A(x, y) = 0\) for every \(x \in [C_A(B), C_A(B)]\) and \(y \in B\). If \(x \in B\) as well, then we get that \(x = 0\), because \(b_A(\cdot, \cdot)\) is nondegenerate on \(B\).

Step (6) of the proof on p37 of [13] states that \(C_A(B)\) is commutative as a Lie algebra over \(k\). Suppose for the sake of a contradiction that

\[(17.4.14) \quad [C_A(B), C_A(B)] \neq \{0\}.\]

This implies that

\[(17.4.15) \quad ([C_A(B), C_A(B)]) \cap Z(C_A(B)) \neq \{0\},\]

because \([C_A(B), C_A(B)]\) is an ideal in \(C_A(B)\), and \(C_A(B)\) is nilpotent as a Lie algebra, as in Section 9.10. Let \(z\) be an element of the left side of (17.4.15) with \(z \neq 0\). Note that \(z \notin B\), by the previous step.

It follows that \(z\) is not ad-diagonalizable as an element of \(A\), as before. Let \(z = z_1 + z_2\) be the abstract Jordan decomposition of \(z\) in \(A\). Thus \(z_2\) is ad-nilpotent as an element of \(A\), \(z_2 \in C_A(B)\), as before, and \(z_2 \neq 0\), because \(z\) is not ad-diagonalizable. Because \(z \in Z(C_A(B))\), \(ad_z\) commutes with \(ad_x\) for every \(x \in C_A(B)\), as in Section 2.4. This implies that \(ad_{z_2}\) commutes with \(ad_x\) for every \(x \in C_A(B)\), as in Section 10.8. Hence

\[(17.4.16) \quad b_A(x, z_2) = tr_A(ad_x \circ ad_{z_2}) = 0\]

for every \(x \in C_A(B)\), as in (17.4.1), because \(ad_{z_2}\) is nilpotent on \(A\). This means that \(z_2 = 0\), because \(b_A(\cdot, \cdot)\) is nondegenerate on \(C_A(B)\).

Step (7) of the proof on p37 of [13] states that (17.4.2) holds. Otherwise, there is an element \(x\) of \(C_A(B)\) not in \(B\), and we can take \(x\) to be ad-nilpotent as an element of \(A\). If \(y \in C_A(B)\), then \([x, y]_A = 0\), by the previous step, so that \(ad_x\) commutes with \(ad_y\) on \(A\), as in Section 2.4. It follows that

\[(17.4.17) \quad b_A(x, y) = tr_A(ad_x \circ ad_y) = 0,\]

as in (17.4.1), because \(ad_x\) is nilpotent on \(A\). This implies that \(x = 0\), because \(b_A(\cdot, \cdot)\) is nondegenerate on \(C_A(B)\).

If \(k = C\), then Lie subalgebras \(B\) of \(A\) with the same types of properties are given in Theorem 3 on p15 of [23].

### 17.5 Self-centralizability and diagonalizability

Let \(k\) be a commutative ring with a multiplicative identity element, and let \((A, [\cdot, \cdot]_A)\) be a Lie algebra over \(k\). Also let \(B\) be a Lie subalgebra of \(A\) that is commutative as a Lie algebra. This implies that

\[(17.5.1) \quad B \subseteq C_A(B),\]
CHAPTER 17. SUBALGEBRAS AND DIAGONALIZABILITY

where $C_A(B)$ is the centralizer of $B$ in $A$, as in Section 7.6. If $B_1$ is another Lie subalgebra of $A$ that is commutative as a Lie algebra, and if $B \subseteq B_1$, then

(17.5.2) $B_1 \subseteq C_A(B_1) \subseteq C_A(B)$.

If

(17.5.3) $C_A(B) = B$,

then it follows that $B_1 \subseteq B$. This means that $B$ is maximal as a commutative Lie subalgebra of $A$ under these conditions. This corresponds to Corollary 1 on p15 of [23].

Conversely, let $x \in C_A(B)$ be given, and let $B(x)$ be the linear span of $B$ and $x$ in $A$, as in (17.4.7). This is a commutative Lie subalgebra of $A$ that contains $B$, as in the previous section. If $B$ is maximal as a commutative Lie subalgebra of $A$, then $x \in B$, and hence (17.5.3) holds.

Suppose now that $k$ is a field, and that $A$ is a finite-dimensional Lie algebra over $k$. Let $B$ be a Lie subalgebra of $A$ such that every element of $B$ is ad-diagonalizable as an element of $A$. This implies that $B$ is commutative as a Lie algebra, as in Section 17.1, so that (17.5.1) holds. If $\alpha$ is an element of the dual $B'$ of $B$, then we let $A_\alpha$ be the set of $x \in A$ such that

(17.5.4) $\text{ad}_w(x) = [w, x]_A = \alpha(w) x$

for every $w \in B$, as in Section 17.2. Let $\Phi_B$ be the set of $\alpha \in B'$ such that $\alpha \neq 0$ and $A_\alpha \neq \{0\}$, as before.

Suppose that $w \in B$ satisfies $\alpha(w) = 0$ for every $\alpha \in \Phi_B$. This implies that $[w, x]_A = 0$ for every $x \in A_\alpha$ when $\alpha \in \Phi_B$, which holds automatically when $\alpha = 0$. It follows that $[w, x]_A = 0$ for every $x \in A$, because $A$ is spanned by the $A_\alpha$’s with $\alpha \in \Phi_B \cup \{0\}$. This means that $w$ is an element of the center $Z(A)$ of $A$ as a Lie algebra. Of course, if $Z(A) = \{0\}$, then $w = 0$. In this case, we get that the linear span of $\Phi_B$ in $B'$ is equal to $B'$. This corresponds to part (a) of the proposition on p37 of [13], and to the statement 2.2 on p45 of [23].

Let $b(\cdot, \cdot)$ be a bilinear form on $A$ that is associative, or equivalently invariant under the adjoint representation on $A$, as in Section 17.3. Suppose that $b(\cdot, \cdot)$ is nondegenerate on $A$, so that the restriction of $b(\cdot, \cdot)$ to $A_0 = C_A(B)$ is nondegenerate, as before. Let us suppose from now on in this section that (17.5.3) holds, which means that the restriction of $b(\cdot, \cdot)$ to $B$ is nondegenerate. If $\alpha \in B'$, then it follows that there is a unique $t_{b,\alpha} \in B$ such that

(17.5.5) $\alpha(w) = b(w, t_{b,\alpha})$

for every $w \in B$.

Let $\alpha \in B'$, $x \in A_\alpha$, and $y \in A_{-\alpha}$ be given, so that

(17.5.6) $[x, y]_A \in A_0 = C_A(B) = B$.

Remember that

(17.5.7) $b(w, [x, y]_A) = \alpha(w) b(x, y)$
17.6. CHARACTERISTIC 0, $Z(A) = \{0\}$

for every $w \in B$, as in Section 17.3. It follows that

\[(17.5.8)\quad [x, y]_A = b(x, y) t_{b, \alpha},\]

because of (17.5.5) and the nondegeneracy of $b(\cdot, \cdot)$ on $B$. This corresponds to part (c) of the proposition on p37 of [13], and to Theorem 3 (iii) on p44 of [23].

Put

\[(17.5.9)\quad B_\alpha = [A_\alpha, A_{-\alpha}]\]

for each $\alpha \in B'$, using the notation in Section 9.2 on the right side. If $\alpha \in \Phi_B$, then there is an $x \in A_\alpha$ with $x \neq 0$, and hence a $y \in A_{-\alpha}$ such that $b(x, y) \neq 0$, as in Section 17.3. This implies that $B_\alpha$ is the same as the one-dimensional linear subspace of $B$ spanned by $t_{b, \alpha}$, because of (17.5.8). This corresponds to part (d) of the proposition on p37 in [23], and to the part of Theorem 2 (b) on p43 of [23] shown in statement 2.3 on p45 of [23].

**17.6 Characteristic 0, $Z(A) = \{0\}$**

Let us continue with the same notation and hypotheses as in the previous section. Let us also suppose in this section that $k$ has characteristic 0, and that $Z(A) = \{0\}$ again. Let $\alpha \in \Phi_B$ be given, and let us show that

\[(17.6.1)\quad \alpha(t_{b, \alpha}) = b(t_{b, \alpha}, t_{b, \alpha}) \neq 0.\]

This corresponds to part (e) of the proposition on p37 of [13], and to the proof of statement 2.4 on p45 of [23].

Suppose for the sake of a contradiction that $\alpha(t_{b, \alpha}) = 0$. This implies that

\[(17.6.2)\quad [t_{b, \alpha}, x]_A = \alpha(t_{b, \alpha}) x = 0\]

for every $x \in A_\alpha$, and that

\[(17.6.3)\quad [t_{b, \alpha}, y]_A = -\alpha(t_{b, \alpha}) y = 0\]

for every $y \in A_{-\alpha}$. As before, there are $x \in A_\alpha$ and $y \in A_{-\alpha}$ such that $b(x, y) \neq 0$, and we can choose them so that $b(x, y) = 1$. Thus

\[(17.6.4)\quad [x, y]_A = t_{b, \alpha},\]

by (17.5.8).

Let $C$ be the linear span of $x$, $y$, and $t_{b, \alpha}$ in $A$. This is a Lie subalgebra of $A$, which is nilpotent and hence solvable as a Lie algebra over $k$, by (17.6.2), (17.6.3), and (17.6.4). The restriction of the adjoint representation on $A$ to $C$ defines a representation of $C$, as a Lie algebra over $k$, on $A$, as a vector space over $k$. Because $t_{b, \alpha} \in [C, C]$, by (17.6.4), we get that $\text{ad}_{t_{b, \alpha}}$ is nilpotent as a linear mapping from $A$ into itself, as in Section 14.14. However, $\text{ad}_{t_{b, \alpha}}$ is diagonalizable as a linear mapping from $A$ into itself, because $t_{b, \alpha} \in B$. It follows that $\text{ad}_{t_{b, \alpha}} = 0$, so that $t_{b, \alpha} \in Z(A)$. This means that $t_{b, \alpha} = 0$, because
Let $\alpha \in \Phi_B$ be given again, and put
\begin{equation}
 h_\alpha = 2 b(t_{b,\alpha}, t_{b,\alpha})^{-1} t_{b,\alpha} = 2 \alpha(t_{b,\alpha})^{-1} t_{b,\alpha}.
\end{equation}

Thus
\begin{equation}
 \alpha(h_\alpha) = 2 \alpha(t_{b,\alpha})^{-1} \alpha(t_{b,\alpha}) = 2.
\end{equation}

Note that $h_\alpha$ is uniquely determined as an element of (17.5.9) by (17.6.6), as in Theorem 2 (b) on p43-4 of [23].

Let $x_\alpha$ be a nonzero element of $A_\alpha$. As before, there is a $y \in A_{-\alpha}$ such that $b(x_\alpha, y) \neq 0$. More precisely, we can choose $y_\alpha \in A_{-\alpha}$ such that
\begin{equation}
 b(x_\alpha, y_\alpha) = 2 b(t_{b,\alpha}, t_{b,\alpha})^{-1}.
\end{equation}

This implies that
\begin{equation}
 [x_\alpha, y_\alpha]_A = h_\alpha,
\end{equation}

by (17.5.8).

Observe that
\begin{equation}
 [h_\alpha, x_\alpha]_A = \alpha(h_\alpha) x_\alpha = 2 x_\alpha,
\end{equation}

using the facts that $h_\alpha \in B$ and $x_\alpha \in A_\alpha$ in the first step. Similarly,
\begin{equation}
 [h_\alpha, y_\alpha]_A = -\alpha(h_\alpha) y_\alpha = -2 y_\alpha,
\end{equation}

because $y_\alpha \in A_{-\alpha}$. This shows that the linear span of $x_\alpha$, $y_\alpha$, and $h_\alpha$ in $A$ is a Lie subalgebra of $A$, which is isomorphic to $sl_2(k)$ as a Lie algebra over $k$. More precisely, $x_\alpha$, $y_\alpha$, and $h_\alpha$ correspond to the usual basis elements $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, and $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ of $sl_2(k)$ under this isomorphism, as in Section 10.2. This corresponds to part (f) of the proposition on p37 of [13], and to part of Theorem 2 (c) on p44 of [23], as in statement 2.5 on p45 of [23].

It is easy to see that
\begin{equation}
 t_{b,-\alpha} = -t_{b,\alpha},
\end{equation}

by the definition (17.5.5) of $t_{b,\alpha}$. Remember that $-\alpha \in \Phi_B$, because $\alpha \in \Phi_B$, as in Section 17.3. Using (17.6.11), we get that
\begin{equation}
 h_{-\alpha} = -h_\alpha,
\end{equation}

as in part (g) of the proposition on p37 of [13].

### 17.7 The dimension of $A_\alpha$

Let us continue with the situation considered in the previous two sections. Thus $k$ is a field of characteristic 0, and $(A, [\cdot, \cdot]_A)$ is a finite-dimensional Lie algebra over $k$. Let $B$ be a Lie subalgebra of $A$ such that every element of $B$ is ad-diagonalizable as an element of $A$, so that $B$ is commutative as a Lie algebra,
17.7. THE DIMENSION OF $A_{\alpha}$

as in Section 17.1. If $\alpha$ is an element of the dual $B'$ of $B$, then $A_{\alpha}$ is the set of $x \in A$ such that
\begin{equation}
\text{ad}_w(x) = [w, x]_A = \alpha(w) x
\end{equation}
for every $w \in B$, as before. In particular, $A_0$ is the same as the centralizer $C_A(B)$ of $B$ in $A$. We suppose here that this is equal to $B$, so that
\begin{equation}
A_0 = C_A(B) = B;
\end{equation}
and that $Z(A) = \{0\}$. Remember that $\Phi_B$ is the set of $\alpha \in B'$ such that $\alpha \neq 0$ and $A_\alpha \neq 0$.

Suppose that $b(\cdot, \cdot)$ is a nondegenerate bilinear form on $A$ that is associative, or equivalently invariant under the adjoint representation on $A$. It follows that $b(\cdot, \cdot)$ is nondegenerate on (17.7.2), as before. If $\alpha \in B'$, then we take $t_{b,\alpha}$ to be the unique element of $B$ such that
\begin{equation}
\alpha(w) = b(w, t_{b,\alpha})
\end{equation}
for every $w \in B$, as in Section 17.5.

Let $\alpha \in \Phi_B$ be given, so that $-\alpha \in \Phi_B$ too, and the restriction of $b(\cdot, \cdot)$ to $A_\alpha + A_{-\alpha}$ is nondegenerate, as in Section 17.3. Remember that the restriction of $b(\cdot, \cdot)$ to each of $A_\alpha$ and $A_{-\alpha}$ is equal to 0. One can use this and the nondegeneracy of $b(\cdot, \cdot)$ on $A_\alpha + A_{-\alpha}$ to get that the dimensions of $A_\alpha$ and $A_{-\alpha}$ are the same, as vector spaces over $k$.

Let $h_\alpha$ be as in (17.6.5), let $x_\alpha$ be a nonzero element of $A_\alpha$, and let $y_\alpha \in A_{-\alpha}$ be as in (17.6.7). Suppose for the sake of a contradiction that the dimension of $A_{\alpha}$ is strictly larger than 1, which means that the dimension of $A_{-\alpha}$ is strictly larger than 1 too. This implies that there is a $y \in A_{-\alpha}$ such that $y \neq 0$ and
\begin{equation}
b(x_\alpha, y) = 0.
\end{equation}
It follows that
\begin{equation}
[x_\alpha, y]_A = 0,
\end{equation}
by (17.5.8). Note that
\begin{equation}
[h_\alpha, y]_A = -\alpha(h_\alpha) y = -2 y,
\end{equation}
because $h_\alpha \in B$, $y \in A_\alpha$, and $\alpha(h_\alpha) = 2$.

Remember that the linear span of $x_\alpha$, $y_\alpha$, and $h_\alpha$ in $A$ is a Lie subalgebra of $A$ that is isomorphic to $sl_2(k)$ as a Lie algebra over $k$. Thus $A$ may be considered as a module over $sl_2(k)$, using the restriction of the adjoint representation on $A$ to this Lie subalgebra of $A$, acting on $A$ as a vector space over $k$. With respect to this representation, $y$ is a maximal or primitive vector of weight $-2$, as in Section 15.2. This contradicts the fact that the weight of $y$ should be nonnegative, as in Section 15.3. This shows that $A_\alpha$ has dimension one as a vector space over $k$, as in Theorem 2 (b) on p43 of [23], and statement 2.6 on p45 of [23].
Of course, this means that the dimension of $A_{-\alpha}$ is one as well. Remember that $B_{\alpha} = [A_{\alpha}, A_{-\alpha}]$ is the same as the one-dimensional linear subspace of $B$ spanned by $h_{\alpha}$, as in the previous two sections. It follows that

$$A_{\alpha} + A_{-\alpha} + B_{\alpha}$$

is the same as the linear span in $A$ of $x_{\alpha}$, $y_{\alpha}$, and $h_{\alpha}$, which is a Lie subalgebra of $A$ isomorphic to $sl_2(k)$, as before. This corresponds to part of Theorem 2 (c) on p44 of [23], and to statement 2.7 on p46 of [23]. Because $A_{-\alpha}$ has dimension one, $y_\alpha$ is uniquely determined by $x_\alpha \in A_\alpha$ and (17.6.7), as in Theorem 2 (c) on p44 of [23] and statement 2.8 on p46 of [23].

Now let us consider the argument given on the bottom of p38 of [13], which can also be used to get some of the same properties as before. Let $C$ be the linear span in $A$ of $B$ and the linear subspaces of the form $A_{c\alpha}$, where $c \in k$ and

$$c \alpha \in \Phi_B,$$

which implies that $c \neq 0$. This is a Lie subalgebra of $A$, because of (17.2.7). In particular, $C$ is a submodule of $A$, as a module over the Lie subalgebra spanned by $x_{\alpha}$, $y_{\alpha}$, and $h_{\alpha}$, because $x_{\alpha}, y_\alpha, h_\alpha \in C$. Thus $C$ may be considered as a module over $sl_2(k)$.

Of course, $[h_\alpha, w]_A = 0$ for every $w \in B$, because $B$ is commutative as a Lie algebra over $k$. If $z \in A_{c\alpha}$ for some $c \in k$, then

$$[h_\alpha, z]_A = c \alpha(h_\alpha) z = 2cz,$$

because $\alpha(h_\alpha) = 2$. This means that the weights of $h_\alpha$ on $C$ consist of 0 and $2c$ for each $c \in k$ satisfying (17.7.8), as in Section 15.1. In particular, the action of $h_\alpha$ on $C$ is diagonalizable in this situation.

Remember that $C$ corresponds to the direct sum of finitely many irreducible submodules, as a module over $sl_2(k)$, by Weyl’s theorem, as in Section 13.1. Any submodule of $C$, as a module over $sl_2(k)$, is mapped into itself by the action of $h_\alpha$, and hence the action of $h_\alpha$ on this submodule is diagonalizable, as in Section 10.6. This implies that any nonzero submodule of $C$ has a maximal or primitive vector, as in Section 15.2. It follows that any nonzero irreducible submodule of $C$ is as in Section 15.3.

The weights of $h_\alpha$ on any nonzero irreducible submodule of $C$ correspond to integers, under the usual embedding of $\mathbb{Q}$ into $k$, as in Section 15.3. This means that the weights of $h_\alpha$ on $C$ correspond to integers. If $c \in k$ satisfies (17.7.8), then we get that $2c$ corresponds to an integer.

If $w \in B$ satisfies $\alpha(w) = 0$, then $[w, x_\alpha]_A = \alpha(w) x_\alpha = 0$ and $[w, y_\alpha]_A = -\alpha(w) y_\alpha = 0$. Of course, $[w, h_\alpha]_A = 0$, because $B$ is commutative as a Lie algebra. Thus the linear span of $x_{\alpha}$, $y_{\alpha}$, and $h_{\alpha}$ acts trivially on the kernel of $\alpha$ in $B$.

Note that the linear span of $x_{\alpha}$, $y_{\alpha}$, and $h_{\alpha}$ is a submodule of $C$. More precisely, this is an irreducible submodule of $C$, because $sl_2(k)$ is simple as a Lie algebra over $k$, as in Section 11.1.
Of course, $B$ is spanned by $h_\alpha$ and the kernel of $\alpha$ in $B$, so that the linear span in $A$ of the kernel of $\alpha$ in $B$, $x_\alpha$, $y_\alpha$, and $h_\alpha$ is the same as the linear span of $B$, $x_\alpha$, and $y_\alpha$. This is a submodule of $C$, and the proof of Weyl’s theorem shows that $C$ corresponds to the direct sum of this submodule and finitely many irreducible submodules of $C$, if necessary.

The elements of $C$ of weight 0 with respect to $h_\alpha$ are in $B$. Thus nonzero elements of irreducible submodules of $C$ complementary to the linear span of $B$, $x_\alpha$, and $y_\alpha$ cannot have weight 0 with respect to $h_\alpha$. This implies that the weights of $h_\alpha$ on these complementary irreducible submodules of $C$ correspond to odd integers, as in Section 15.3. It follows that the only even integers which can correspond to weights of $h_\alpha$ on $C$ are 0 and \pm 2, which are the weights of $h_\alpha$ on $B$, $x_\alpha$, and $y_\alpha$.

In particular, 4 does not correspond to a weight of $h_\alpha$ on $C$, which implies that
\[
2 \alpha \notin \Phi_B.
\]
If $\alpha/2$ were in $\Phi_B$, then we could apply the same argument to it, to get that $\alpha \notin \Phi_B$. Thus
\[
\alpha/2 \notin \Phi_B.
\]
This shows that 1 is not a weight of $h_\alpha$ on $C$.

Suppose for the sake of a contradiction that $C_1$ is a nonzero irreducible submodule of $C$ complementary to the linear span of $B$, $x_\alpha$, and $y_\alpha$. Thus $C_1$ does not contain any nonzero elements with weight 0 or 1 with respect to $h_\alpha$, as in the previous two paragraphs. This is a contradiction, as in Section 15.3. This means that $C$ is the same as the linear span of $B$, $x_\alpha$, and $y_\alpha$.

It follows in particular that $A_\alpha$ and $A_{-\alpha}$ are spanned by $x_\alpha$ and $y_\alpha$, respectively, so that the dimensions of $A_\alpha$ and $A_{-\alpha}$ are both equal to one, as vector spaces over $k$. We also get that $\pm \alpha$ are the only multiples of $\alpha$ by elements of $k$ in $\Phi_B$. This corresponds to parts (a) and (b) of the proposition on p.39 of [13]. The second statement corresponds to the second part of Theorem 2 (a) on p.43 of [23] as well.

### 17.8 Other elements of $\Phi_B$

Let us continue with the same notation and hypotheses as in the previous section. Let $\alpha, \beta \in \Phi_B$ be given, and remember that $h_\alpha \in B$ is as in (17.6.5). Let $x_\alpha$ be a nonzero element of $A_\alpha$ again, and let $y_\alpha \in A_{-\alpha}$ be as in (17.6.7), so that the linear span of $x_\alpha$, $y_\alpha$, and $h_\alpha$ in $A$ is a Lie subalgebra of $A$ that is isomorphic to $sl_2(k)$, as a Lie algebra over $k$. Thus $A$ may be considered as a module over the linear span of $x_\alpha$, $y_\alpha$, and $h_\alpha$, or over $sl_2(k)$. Weyl’s theorem implies that $A$ corresponds to the direct sum of finitely many irreducible submodules. Any submodule of $A$ is mapped into itself by $ad_{h_\alpha}$, and $ad_{h_\alpha}$ is diagonalizable on the submodule. This implies that any nonzero submodule of $A$ has a maximal or primitive vector, as in Section 15.2, and hence a nonzero irreducible submodule
is as in Section 15.3. In particular, the eigenvalues of $\text{ad}_{h_{\alpha}}$ on $A$ correspond to integers, with respect to the usual embedding of $\mathbb{Q}$ in $k$.

Let $y$ be a nonzero element of $A_{\beta}$, so that

\begin{equation}
\text{ad}_{h_{\alpha}}(y) = [h_{\alpha}, y]_A = \beta(h_{\alpha}) y.
\end{equation}

Thus $\beta(h_{\alpha})$ corresponds to an integer, with respect to the standard embedding of $\mathbb{Q}$ in $k$, as in the preceding paragraph. This is part of statement 2.9 on p46 of [23].

Let $n$ be the integer corresponding to $\beta(h_{\alpha})$, and put

\begin{equation}
z = (\text{ad}_{y_{\alpha}})^n(y) \quad \text{when } n \geq 0 \quad \text{and} \quad z = (\text{ad}_{x_{\alpha}})^{-n}(y) \quad \text{when } n \leq 0.
\end{equation}

One can check that $z \neq 0$, because $A$ corresponds to the direct sum of finitely many irreducible submodules, each of which is as in Section 15.3. We also have that

\begin{equation}
z \in A_{\beta - \beta(h_{\alpha}) \alpha},
\end{equation}

because $y \in A_{\beta}$, $x_{\alpha} \in A_{\alpha}$, and $y_{\alpha} \in A_{-\alpha}$, and using (17.2.7). In particular,

\begin{equation}
A_{\beta - \beta(h_{\alpha}) \alpha} \neq \{0\},
\end{equation}

because $z \neq 0$.

Let us check that

\begin{equation}
\beta - \beta(h_{\alpha}) \alpha \neq 0.
\end{equation}

Otherwise, if $\beta - \beta(h_{\alpha}) \alpha = 0$, then

\begin{equation}
0 = \beta(h_{\alpha}) - \beta(h_{\alpha}) \alpha(h_{\alpha}) = \beta(h_{\alpha}) - 2 \beta(h_{\alpha}) = -\beta(h_{\alpha}),
\end{equation}

because $\alpha(h_{\alpha}) = 2$. In this case, we get that $\beta = 0$, contradicting the hypothesis that $\beta \in \Phi_B$. Thus (17.8.5) holds, and hence

\begin{equation}
\beta - \beta(h_{\alpha}) \alpha \in \Phi_B,
\end{equation}

by (17.8.4). This corresponds to statement 2.9 on p46 of [23].

If $\beta = \alpha$, then

\begin{equation}
\beta - \beta(h_{\alpha}) \alpha = \alpha - \alpha(h_{\alpha}) \alpha = \alpha - 2 \alpha = -\alpha,
\end{equation}

because $\alpha(h_{\alpha}) = 2$, as before. Similarly, if $\beta = -\alpha$, then

\begin{equation}
\beta - \beta(h_{\alpha}) \alpha = -\alpha - (-\alpha(h_{\alpha})) \alpha = -\alpha + 2 \alpha = \alpha.
\end{equation}

This is part of statement 2.10 on p46 of [23].

Here is another proof of the fact that $c \in k$ satisfies $c \alpha \in \Phi_B$ only when $c = \pm 1$, as in statement 2.11 on p4 of [23]. If $\beta = c \alpha \in \Phi_B$, then $\beta(h_{\alpha}) = c \alpha(h_{\alpha}) = 2c$ corresponds to an integer, as before. Similarly, we can interchange
the roles of \( \alpha \) and \( \beta \), to get that \( 2/c \) corresponds to an integer. Thus the only possibilities for \( c \) are \( \pm 1, \pm 1/2, \) and \( \pm 2 \).

It suffices to show that \( 2 \alpha \not\in \Phi_B \), which is to say that \( c \neq 2 \). This will imply that \( \alpha/2 \not\in \Phi_B \), since otherwise we would have that \( \alpha \not\in \Phi_B \). Similarly, this will show that \( c \neq -1/2, -2 \), because \(-\alpha \not\in \Phi_B \), as in Section 17.3.

Suppose for the sake of a contradiction that \( 2 \alpha \in \Phi_B \), and let \( y \) be a nonzero element of \( A_{2\alpha} \). Thus

\[
[h_\alpha, y]_A = 2 \alpha(h_\alpha) y = 4y.
\]

(17.8.10)

Note that \( \text{ad}_{x_\alpha}(y) \in A_{3\alpha} \), by (17.2.7), and because \( x_\alpha \in A_\alpha \). However, \( A_{3\alpha} = \{0\} \), because \( 3 \alpha \not\in \Phi_B \), as before. This means that

\[
\text{ad}_{x_\alpha}(y) = 0.
\]

(17.8.11)

Observe that

\[
\text{ad}_{h_\alpha}(y) = \text{ad}_{x_\alpha}(\text{ad}_{y_\alpha}(y)) - \text{ad}_{y_\alpha}(\text{ad}_{x_\alpha}(y)) = \text{ad}_{x_\alpha}(\text{ad}_{y_\alpha}(y)),
\]

(17.8.12)

using (17.6.8) in the first step, and (17.8.11) in the second step. Using (17.2.7) again, we get that \( \text{ad}_{y_\alpha}(y) \in A_\alpha \), because \( y_\alpha \in A_{-\alpha} \) and \( y \in A_{2\alpha} \). This means that \( \text{ad}_{y_\alpha}(y) \) is a multiple of \( x_\alpha \), because \( A_\alpha \) has dimension one as a vector space over \( k \), as in the previous section. It follows that

\[
\text{ad}_{x_\alpha}(\text{ad}_{y_\alpha}(y)) = 0,
\]

(17.8.13)

because \( \text{ad}_{x_\alpha}(x_\alpha) = 0 \). This contradicts (17.8.10), because of (17.8.12) and the fact that \( y \neq 0 \).

17.9 Adding elements of \( \Phi_B \)

Let us continue with the same notation and hypotheses as in the previous two sections. Let \( \alpha, \beta \in \Phi_B \) be given again, and suppose that

\[
\beta \neq \pm \alpha.
\]

(17.9.1)

Thus, for each \( j \in \mathbb{Z} \),

\[
\beta + j\alpha \neq 0,
\]

(17.9.2)

as in the previous sections. Let \( E \) be the linear span in \( A \) of the subspaces

\[
A_{\beta + j\alpha},
\]

(17.9.3)

where \( j \in \mathbb{Z} \). More precisely, we may restrict our attention to \( j \in \mathbb{Z} \) such that

\[
\beta + j\alpha \in \Phi_B,
\]

(17.9.4)

since otherwise (17.9.3) is equal to \( \{0\} \).
Remember that $h_\alpha \in B$ is as in (17.6.5), and let $x_\alpha$ be a nonzero element of $A_\alpha$, again. Also let $y_\alpha \in A_{-\alpha}$ be as in (17.6.7), so that the linear span of $x_\alpha$, $y_\alpha$, and $h_\alpha$ in $A$ is a Lie subalgebra of $A$ that is isomorphic to $sl_2(k)$, as a Lie algebra over $k$. As before, $A$ may be considered as a module over the linear span of $x_\alpha$, $y_\alpha$, and $h_\alpha$, or over $sl_2(k)$. In fact, $E$ is a submodule of $A$, because of (17.2.7). The weights of $h_\alpha$ on $E$ are given by

\[(17.9.5) \quad \beta(h_\alpha) + j \alpha(h_\alpha)\]

for $j \in \mathbb{Z}$ such that (17.9.4) holds, as in Section 15.1. Of course, (17.9.5) is the same as

\[(17.9.6) \quad \beta(h_\alpha) + 2j,\]

because $\alpha(h_\alpha) = 2$. Note that $j = 0$ satisfies (17.9.4) automatically.

Weyl’s theorem implies that $E$ corresponds to the direct sum of finitely many irreducible submodules, as in Section 13.1. Any submodule of $E$ is mapped into itself by the action of $h_\alpha$, and the action of $h_\alpha$ on this submodule is diagonalizable, as in Section 10.6. Hence any nonzero submodule of $E$ has a maximal or primitive vector, as in Section 15.2. This implies that any nonzero irreducible submodule of $E$ is as in Section 15.3.

Remember that (17.9.3) has dimension one as a vector space over $k$ when (17.9.4) holds, as in Section 17.7. One can use this to check that $E$ is irreducible as a module over $sl_2(k)$. More precisely, note that at most one of 0 and 1 can be of the form (17.9.6), as mentioned near the top of p39 of [13].

Thus there is a nonnegative integer $m$ such that $E$ is as in Section 15.3, as a module over $sl_2(k)$. In particular, the dimension of $E$ is $m + 1$, as a vector space over $k$.

Let $q$ and $r$ be the largest integers such that (17.9.4) holds with $j = q$ and $j = -r$, respectively. Note that $q, r \geq 0$, because (17.9.4) holds when $j = 0$. The maximal and minimal weights of $h_\alpha$ on $E$ are

\[(17.9.7) \quad m = \beta(h_\alpha) + 2q\]

and

\[(17.9.8) \quad -m = \beta(h_\alpha) - 2r,\]

as in Section 15.3. In particular,

\[(17.9.9) \quad \beta(h_\alpha) = r - q.\]

Suppose that $j \in \mathbb{Z}$ satisfies

\[(17.9.10) \quad -r \leq j \leq q,\]

so that

\[(17.9.11) \quad \beta(h_\alpha) - 2r \leq \beta(h_\alpha) + 2j \leq \beta(h_\alpha) + 2q.\]

Under these conditions, $\beta(h_\alpha) + 2j$ is a weight of $h_\alpha$ on $E$, as in Section 15.3. This means that (17.9.4) holds, as in part (e) of the proposition on p39 of [13].
Similarly, if \( j \in \mathbb{Z} \) satisfies
\[
- r \leq j \leq q - 1,
\]
then the restriction of \( \text{ad}_{x_\alpha} \) to \( A_{\beta + j \alpha} \) is a one-to-one mapping onto
\[
A_{\beta + (j+1) \alpha}.
\]
This follows from Section 15.3 again. These properties of \( E, q, \) and \( r \) correspond to statement 2.12 on p46 of [23].

The same type of arguments are discussed starting at the bottom of p38 and continuing on p39 of [13], as before. In particular, these arguments can also be used to get that \( \beta(h_\alpha) \) corresponds to an integer for which (17.8.7) holds, as in part (c) of the proposition on p39 of [13].

Suppose now that \( \alpha + \beta \in \Phi_B \). This means that (17.9.4) holds with \( j = 1 \), and in particular that \( q \geq 1 \). Thus \( j = 0 \) satisfies (17.9.12), so that \( \text{ad}_{x_\alpha} \) maps \( A_\beta \) onto \( A_{\beta + \alpha} \), as before. This implies that
\[
[A_\alpha, A_\beta] = A_{\alpha + \beta},
\]
as in Theorem 2 (d) on p44 of [23], and statement 2.13 on p47 of [23]. This corresponds to part (d) of the proposition on p39 of [13] as well.

17.10 The dual of \( B \)

Let \( k \) be a field of characteristic 0 again, and let \( (A, [\cdot, \cdot]_A) \) be a finite-dimensional Lie algebra over \( k \). Also let \( B \) be a Lie subalgebra of \( A \) such that every element of \( B \) is ad-diagonalizable as an element of \( A \), so that \( B \) is commutative as a Lie algebra, as in Section 17.1. If \( \alpha \) is an element of the dual \( B' \) of \( B \), then put
\[
A_\alpha = \{ x \in A : \text{ for every } w \in B, \text{ ad}_w(x) = [w, x]_A = \alpha(w) x \},
\]
as before. Thus \( A_0 \) is the same as the centralizer \( C_A(B) \) of \( B \) in \( A \), and we suppose again that this is equal to \( B \), so that
\[
A_0 = C_A(B) = B.
\]
We suppose that the center \( Z(A) \) of \( A \) as a Lie algebra is equal to \( \{0\} \) again too.

Suppose that \( b(\cdot, \cdot) \) is a nondegenerate bilinear form on \( A \) that is associative, or equivalently invariant under the adjoint representation on \( A \), as before. This implies that that the restriction of \( b(\cdot, \cdot) \) to (17.10.2) is nondegenerate, as in Section 17.3. If \( \alpha \in B' \), then there is a unique \( t_{\alpha, \alpha} \in B \) such that
\[
\alpha(x) = b(x, t_{\alpha, \alpha})
\]
for every \( w \in B \), as in Section 17.5. Note that \( \alpha \mapsto t_{\alpha, \alpha} \) is a one-to-one linear mapping from \( B' \) onto \( B \).
If \( \alpha, \beta \in B' \), then put
\[
b'(\alpha, \beta) = b(t_{b, \alpha}, t_{b, \beta}).
\]
This defines a bilinear form on \( B' \). Equivalently,
\[
b'(\alpha, \beta) = \beta(t_{b, \alpha})
\]
for every \( \alpha, \beta \in B' \). Observe that (17.10.4) is nondegenerate on \( B' \).

Let \( \Phi_B \) be the set of \( \alpha \in B' \) such that \( \alpha \neq 0 \) and \( A_\alpha \neq \{0\} \), as before. If \( \alpha \in \Phi_B \), then
\[
b'(\alpha, \alpha) = b(t_{b, \alpha}, t_{b, \alpha}) \neq 0,
\]
as in Section 17.6. In this case, we put
\[
h_\alpha = 2b(t_{b, \alpha}, t_{b, \alpha})^{-1}t_{b, \alpha} = 2b'(\alpha, \alpha)^{-1}t_{b, \alpha},
\]
as before, so that \( \alpha(h_\alpha) = 2 \). If \( \beta \in B' \), then
\[
2b'(\alpha, \beta)b'(\alpha, \alpha)^{-1} = 2\beta(t_{b, \alpha})b'(\alpha, \alpha)^{-1} = \beta(h_\alpha).
\]
If \( \beta \in \Phi_B \), then (17.10.8) corresponds to an integer, with respect to the standard embedding of \( \mathbb{Q} \) into \( k \), as in Section 17.8.

Remember that the linear span of \( \Phi_B \) in \( B' \) is equal to \( B' \), as in Section 17.5. Let \( \alpha_1, \ldots, \alpha_n \) be a basis for \( B' \) consisting of elements of \( \Phi_B \). Let \( \beta \in \Phi_B \) be given, so that \( \beta \) can be expressed in a unique way as
\[
\beta = \sum_{j=1}^{n} c_j \alpha_j,
\]
where \( c_j \in k \) for each \( j = 1, \ldots, n \). We would like to show that \( c_j \) corresponds to an element of \( \mathbb{Q} \) for every \( j = 1, \ldots, n \), with respect to the standard embedding of \( \mathbb{Q} \) into \( k \), as discussed near the bottom of p39 of [13].

Of course,
\[
b'(\alpha_l, \beta) = \sum_{j=1}^{n} c_j b'(\alpha_l, \alpha_j)
\]
for each \( l = 1, \ldots, n \). This implies that
\[
2b'(\alpha_l, \beta)b'(\alpha_l, \alpha_l)^{-1} = \sum_{j=1}^{n} 2b'(\alpha_l, \alpha_j)b'(\alpha_l, \alpha_l)^{-1} c_j
\]
for every \( l = 1, \ldots, n \). The left side corresponds to an integer for every \( l = 1, \ldots, n \), because \( \beta \in \Phi_B \), as before. Similarly,
\[
2b'(\alpha_l, \alpha_j)b'(\alpha_l, \alpha_l)^{-1}
\]
corresponds to an integer for every \( j, l = 1, \ldots, n \).
17.11. USING THE KILLING FORM

Note that \((b'(\alpha,\alpha))\) is invertible as an \(n \times n\) matrix with entries in \(k\), because \(b'(\cdot,\cdot)\) is nondegenerate as a bilinear form on \(B'\). This implies that the \(n \times n\) matrix with entries (17.10.12) is invertible as well. Equivalently, the determinant of this matrix is not 0. As in the preceding paragraph, this corresponds to an \(n \times n\) matrix of integers, whose determinant is not 0.

Thus the entries of the inverse of the matrix in (17.10.12) correspond to rational numbers. This permits us to solve (17.10.11), to get that \(c_j\) corresponds to an element of \(Q\) for each \(j = 1, \ldots, n\), as desired.

17.11 Using the Killing form

Let us continue with the same notation and hypotheses as in the previous section, with some additional conditions, as follows. Suppose that \(A\) is semisimple, and let

\[ b_A(x,y) = \text{tr}_A(\text{ad}_x \circ \text{ad}_y) \]  

(17.11.1)

be the Killing form on \(A\). Thus \(b_A\) is associative on \(A\), or equivalently invariant under the adjoint representation on \(A\), and \(b_A\) is nondegenerate on \(A\), as in Sections 7.9 and 10.13. In particular, we can take \(b = b_A\) in the previous section.

Remember that the restriction of \(b_A\) to (17.10.2) is nondegenerate, as before. If \(\alpha \in B'\), then there is a unique \(t_{\alpha} = t_{b_A,\alpha} \in B\) such that

\[ \alpha(w) = b_A(w,t_{\alpha}) \]  

(17.11.2)

for every \(w \in B\), as in (17.10.3). Put

\[ b'_A(\alpha,\beta) = b_A(t_{\alpha},t_{\beta}) \]  

(17.11.3)

for every \(\alpha, \beta \in B'\), as in (17.10.4). Equivalently,

\[ b'_A(\alpha,\beta) = \beta(t_{\alpha}) \]  

(17.11.4)

for every \(\alpha, \beta \in B'\), as in (17.10.5).

If \(x,y \in B\), then

\[ b_A(x,y) = \sum_{\alpha \in \Phi_B} \alpha(x) \alpha(y). \]  

(17.11.5)

To see this, remember that \(A\) corresponds to the direct sum of the subspaces \(A_\alpha, \alpha \in \Phi_B \cup \{0\}\), as a vector space over \(k\), as in Section 17.2. These subspaces are mapped into themselves by \(\text{ad}_x\) and \(\text{ad}_y\), so that the trace of \(\text{ad}_x \circ \text{ad}_y\) over \(A\) is the same as the sum of the traces over \(A_\alpha, \alpha \in \Phi_B \cup \{0\}\). The trace over \(A_0\) is equal to 0, because \(\text{ad}_x = \text{ad}_y = 0\) on \(A_0\), and so it suffices to consider the sum over \(\alpha \in \Phi_B\). By construction, \(\text{ad}_x\) and \(\text{ad}_y\) correspond to multiplication by \(\alpha(x)\) and \(\alpha(y)\) on \(A_\alpha\). If \(\alpha \in \Phi_B\), then the dimension of \(A_\alpha\) is equal to one, as in Section 17.7. This implies (17.11.5).
CHAPTER 17. SUBALGEBRAS AND DIAGONALIZABILITY

If $\lambda, \mu \in B'$, then we get that

$$(17.11.6) \quad b'_A(\lambda, \mu) = b_A(t_\lambda, t_\mu) = \sum_{\alpha \in \Phi_B} \alpha(t_\lambda) \alpha(t_\mu).$$

It follows that

$$(17.11.7) \quad b'_A(\lambda, \mu) = \sum_{\alpha \in \Phi_B} b'_A(\lambda, \alpha) b'_A(\mu, \alpha),$$

by (17.11.4). In particular,

$$(17.11.8) \quad b'_A(\lambda, \lambda) = \sum_{\alpha \in \Phi_B} b'_A(\lambda, \alpha)^2.$$

Suppose that $\beta \in \Phi$, so that $b'_A(\beta, \beta) \neq 0$, as in (17.10.6). Using (17.11.8) with $\lambda = \beta$ and multiplying both sides by $1/b'_A(\beta, \beta)^2$, we get that

$$(17.11.9) \quad 1/b'_A(\beta, \beta) = \sum_{\alpha \in \Phi_B} b'_A(\beta, \alpha)^2/b'_A(\beta, \beta)^2.$$

If $\alpha \in \Phi_B$, then

$$(17.11.10) \quad 2b'_A(\beta, \alpha)/b'_A(\beta, \beta)$$

corresponds to an integer under the standard embedding of $Q$ into $k$, as mentioned in the previous section. This implies that the right side of (17.11.9) corresponds to a rational number. Thus $b'_A(\beta, \beta)$ corresponds to a rational number. It follows that

$$(17.11.11) \quad b'_A(\beta, \alpha)$$

corresponds to a rational number as well when $\alpha \in \Phi_B$, because (17.11.10) corresponds to an integer. This corresponds to some of the remarks near the top of p40 of [13].

17.12 Vector spaces over $Q$

Let us return for the moment to the situation considered in Section 17.10. We may consider the dual $B'$ of $B$ as a vector space over the rational numbers, using the standard embedding of $Q$ into $k$, as mentioned in the previous section. This implies that the right side of (17.11.9) corresponds to a rational number. Thus $b'_A(\beta, \beta)$ corresponds to a rational number. It follows that

$$(17.11.11) \quad b'_A(\beta, \alpha)$$

corresponds to a rational number as well when $\alpha \in \Phi_B$, because (17.11.10) corresponds to an integer. This corresponds to some of the remarks near the top of p40 of [13].
17.13. VECTOR SPACES OVER $\mathbb{R}$

with respect to the standard embedding of $\mathbb{Q}$ into $k$, as before. It follows that $b'_A(\alpha, \beta)$ corresponds to a rational number when $\alpha, \beta \in E_Q$. Let $(\alpha, \beta)_{E_Q}$ be the rational number that corresponds to $b'_A(\alpha, \beta)$ when $\alpha, \beta \in E_Q$. This defines a bilinear form on $E_Q$, as a vector space over $\mathbb{Q}$. If $\lambda, \mu \in E_Q$, then

\[(\lambda, \mu)_{E_Q} = \sum_{\alpha \in \Phi_B} (\lambda, \alpha)_{E_Q} (\mu, \alpha)_{E_Q}, \tag{17.12.1}\]

as in (17.11.7).

It follows that

\[(\lambda, \lambda)_{E_Q} = \sum_{\alpha \in \Phi_B} (\lambda, \alpha)^2_{E_Q} \tag{17.12.2}\]

for every $\lambda \in E_Q$. This implies that

\[(\lambda, \lambda)_{E_Q} \geq 0 \tag{17.12.3}\]

for every $\lambda \in E_Q$. More precisely, $(\lambda, \lambda)_{E_Q} = 0$ if and only if

\[(\lambda, \alpha)_{E_Q} = 0 \text{ for every } \alpha \in \Phi_B. \tag{17.12.4}\]

Let us check that this can only happen when $\lambda = 0$.

Of course, (17.12.4) is the same as saying that

\[b'_A(\lambda, \alpha) = 0 \tag{17.12.5}\]

for every $\alpha \in \Phi_B$. This implies that (17.12.5) holds for every $\alpha \in B'$, because the linear span of $\Phi_B$ in $B'$, as a vector space over $k$, is equal to $B'$. It follows that $\lambda = 0$, because $b'_A$ is nondegenerate on $B'$. This corresponds to some more of the remarks near the top of p40 in [13].

17.13 Vector spaces over $\mathbb{R}$

Let us return again for the moment to the situation considered in Section 17.10, and the beginning of the previous section. Thus $E_Q$ is the linear subspace of $B'$, as a vector space over $\mathbb{Q}$, spanned by $\Phi_B$. Let $n$ be the dimension of $B$ and $B'$ as vector spaces over $k$, which is the same as the dimension of $E_Q$ as a vector space over $\mathbb{Q}$. Using any basis for $E_Q$, we get that $E_Q$ is isomorphic to $\mathbb{Q}^n$, as a vector space over $\mathbb{Q}$. We can get a vector space $E_R$ over the real numbers with the same basis, using formal linear combinations with coefficients in $\mathbb{R}$. This amounts to taking $\mathbb{R}^n$ as a vector space over $\mathbb{R}$ that contains $\mathbb{Q}^n$. This does not depend on the choice of basis for $E_Q$, up to suitable equivalence.

In particular, $E_Q$ corresponds to a subset of $E_R$.

Suppose now, for the rest of the section, that we are in the more particular situation considered in Section 17.11 again, and as in the previous section. Thus $(\alpha, \beta)_{E_Q}$ is defined as a bilinear form on $E_Q$, as a vector space over $\mathbb{Q}$, as before.
This has a natural extension to a bilinear form \((\alpha, \beta)_{E_R}\) on \(E_R\), as a vector space over \(R\). If \(\lambda, \mu \in E_R\), then

\[(\lambda, \mu)_{E_R} = \sum_{\alpha \in \Phi_B} (\lambda, \alpha)_{E_R} (\mu, \alpha)_{E_R},\]

as in (17.12.1). This implies that

\[(\lambda, \lambda)_{E_R} = \sum_{\alpha \in \Phi_B} (\lambda, \alpha)^2_{E_R} \geq 0\]

for every \(\lambda \in E_R\).

Let us verify that \(\lambda \in E_R\) satisfies \((\lambda, \lambda)_{E_R} = 0\) only when \(\lambda = 0\), so that \((\cdot, \cdot)_{E_R}\) defines an inner product on \(E_R\). One way to do this is to use an orthogonal basis for \(E_Q\) with respect to \((\cdot, \cdot)_{E_Q}\), which is an orthogonal basis for \(E_R\) with respect to \((\cdot, \cdot)_{E_R}\) as well. If \(\lambda \in E_R\) and \(\lambda \neq 0\), then one can get that \((\lambda, \lambda)_{E_R} > 0\) from the analogous condition for the basis vectors in \(E_Q\).

Alternatively, let \(\alpha_1, \ldots, \alpha_n\) be elements of \(\Phi_B\) that form a basis for \(B'\) as a vector space over \(k\), so that \(\alpha_1, \ldots, \alpha_n\) forms a basis for \(E_Q\) as a vector space over \(Q\) too, as before. If \(\mu \in E_Q\) satisfies

\[(\mu, \alpha_j)_{E_Q} = 0\]

for each \(j = 1, \ldots, n\), then

\[b'_A(\mu, \alpha_j) = 0\]

for every \(j = 1, \ldots, n\). This implies that

\[b'_A(\mu, \alpha) = 0\]

for every \(\alpha \in B'\), and hence that \(\mu = 0\), because \(b'_A\) is nondegenerate on \(B'\). Of course,

\[\mu \mapsto (\mu, \alpha_j)_{E_Q}\]

defines a linear functional on \(E_Q\) for each \(j = 1, \ldots, n\). Using these linear functionals, we get a mapping from \(E_Q\) into \(Q^n\) that is linear over \(Q\). The kernel of this linear mapping is trivial, by the previous remarks. It follows that this linear mapping sends \(E_Q\) onto \(Q^n\), because \(E_Q\) has dimension \(n\) as a vector space over \(Q\).

Similarly,

\[\mu \mapsto (\mu, \alpha_j)_{E_R}\]

is a linear functional on \(E_R\) for each \(j = 1, \ldots, n\). Using these linear functionals, we get a mapping from \(E_R\) into \(R^n\) that is linear over \(R\). The image of this linear mapping contains \(Q^n\), as in the preceding paragraph. This implies that this linear mapping sends \(E_R\) onto \(R^n\), because the image is a linear subspace of \(R^n\), as a vector space over \(R\). It follows that this linear mapping is injective, because \(E_R\) has dimension \(n\), as a vector space over \(R\). Thus, if \(\lambda \in E_R\) satisfies

\[(\lambda, \alpha_j)_{E_R} = 0\]

for every \(j = 1, \ldots, n\), then \(\lambda = 0\). If \((\lambda, \lambda)_{E_R} = 0\), then (17.13.2) holds for each \(j = 1, \ldots, n\), by (17.13.2), so that \(\lambda = 0\). This corresponds to some of the remarks near the top of p40 of [13] again.
17.14 Normalizers and centralizers

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. Also let $B$ be a Lie subalgebra of $A$. Remember that the centralizer $C_A(B)$ of $B$ in $A$ consists of the $x \in A$ such that $[x, w]_A = 0$ for every $w \in B$, as in Section 7.6. The normalizer $N_A(B)$ of $B$ in $A$ is the set of $x \in A$ such that $[x, w]_A \in B$ for every $w \in B$, as in Section 9.8. Thus

$$B \subseteq N_A(B),$$

because $B$ is a Lie subalgebra of $A$. Note that $C_A(B) = N_A(B)$.

Both $C_A(B)$ and $N_A(B)$ are Lie subalgebras of $A$, as before. Of course, if $B$ is commutative as a Lie algebra over $k$, then $B = C_A(B)$.

Suppose from now on in this section that $k$ is a field, and that $A$ is a finite-dimensional Lie algebra over $k$. Let $B$ be a finite-dimensional Lie subalgebra of $A$ such that every element of $B$ is ad-diagonalizable as an element of $A$. Thus $B$ is commutative as a Lie algebra over $k$, as in Section 17.1. Let $B'$ be the dual of $B$, as a vector space over $k$, and for each $w \in B'$, let $A_\alpha$ be the set of $x \in A$ such that $\text{ad}_w(x) = [w, x]_A = \alpha(w) x$ for every $w \in B$, as usual. If $\alpha = 0$, then this is the same as $C_A(B)$, as before.

In particular, $B \subseteq C_A(B) = A_0$. Let us suppose that $B \neq \{0\}$, to avoid trivialities, so that $A_0 \neq \{0\}$. Let $\Phi_B$ be the set of $\alpha \in B'$ such that $\alpha \neq 0$ and $A_\alpha \neq \{0\}$ again. Remember that $A$ corresponds to the direct sum of $A_\alpha$, $\alpha \in \Phi_B \cup \{0\}$, as a vector space over $k$, as in Section 17.2.

Let us check that

$$A_0 = N_A(B)$$

under these conditions. Of course, $A_0 \subseteq N_A(B)$, as in (17.14.2). Let $x \in A$ be given, so that $x$ can be expressed in a unique way as

$$x = \sum_{\alpha \in \Phi_B \cup \{0\}} x_\alpha,$$

where $x_\alpha \in A_\alpha$ for each $\alpha \in \Phi_B \cup \{0\}$. If $w \in B$, then

$$[w, x]_A = \sum_{\alpha \in \Phi_B \cup \{0\}} [w, x_\alpha]_A = \sum_{\alpha \in \Phi_B} \alpha(w) x_\alpha.$$

This is in $B$, or even $A_0$, only when it is equal to 0. If this happens for every $w \in B$, then it follows that $x \in A_0 = C_A(B)$. Equivalently, this means that $x_\alpha = 0$ for every $\alpha \in \Phi_B$. This shows that $N_A(B) \subseteq A_0$, so that (17.14.3) holds.

If $C_A(B) = B$, then we get that $N_A(B) = B$. This corresponds to Exercise 5 on p40 of [13].
17.15 Semisimplicity and diagonalizability

Let $k$ be a field, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$ with positive finite dimension, as a vector space over $k$. Also let $B$ be a Lie subalgebra of $A$ such that every element of $B$ is ad-diagonalizable as an element of $A$. Thus $B$ is commutative as a Lie algebra over $k$, as in Section 17.1. Let $B'$ be the dual of $B$, as a vector space over $k$, and for each $\alpha \in B'$, let $A_{\alpha}$ be the set of $x \in A$ such that $\text{ad}_w(x) = [w, x]_A = \alpha(w) x$ for every $w \in B$, as before. Remember that $A_0$ is the same as the centralizer $C_A(B)$ of $B$ in $A$, which contains $B$, because $B$ is commutative as a Lie algebra.

Suppose that $B = C_A(B) = A_0$.

This implies that $B \neq \{0\}$, because $A \neq \{0\}$, by hypothesis. Let $\Phi_B$ be the set of $\alpha \in B'$ such that $\alpha \neq 0$ and $A_{\alpha} \neq \{0\}$ again, so that $A$ corresponds to the direct sum of $A_{\alpha}$, $\alpha \in \Phi_B \cup \{0\}$, as a vector space over $k$, as in Section 17.2. Note that the center $Z(A)$ of $A$ as a Lie algebra is contained in $A_0$. Let us suppose also that $Z(A) = \{0\}$.

Suppose in addition that for each $\alpha \in \Phi_B$ and $x_\alpha \in A_{\alpha}$ with $x_\alpha \neq 0$ there is a $y_\alpha \in A_{-\alpha}$ with the following property. Put

$$ z_\alpha = [x_\alpha, y_\alpha]_A, $$

so that $z_\alpha \in A_0$, as in Section 17.2. This means that $z_\alpha \in B$, by (17.15.1), and we ask that

$$ \alpha(z_\alpha) \neq 0. $$

We have seen that this holds under some additional conditions, as in Section 17.6.

We would like to show that $A$ is semisimple as a Lie algebra over $k$ in this situation. This corresponds to Step (13) on p100 of [13], and to (k) on p55 of [23]. Let $C$ be an ideal in $A$ such that $C$ is commutative as a Lie algebra. It suffices to show that $C = \{0\}$, as in Section 9.4.

If $w \in B$, then $\text{ad}_w$ maps $C$ into itself, because $C$ is an ideal in $A$. The restriction of $\text{ad}_w$ to $C$ is diagonalizable, as in Section 10.6, because $\text{ad}_w$ is diagonalizable on $A$, by hypothesis. Remember that the mappings $\text{ad}_w$, with $w \in B$, commute on $A$, because $B$ is commutative as a Lie algebra, as in Section 17.2. This implies that the linear mappings $\text{ad}_w$, $w \in B$, are simultaneously diagonalizable on $C$, as before. If $C \neq \{0\}$, then it follows that $C$ is spanned by its intersections with the subspaces $A_{\alpha}$, with $\alpha \in \Phi_B \cup \{0\}$.

Suppose for the sake of a contradiction that there is an $\alpha \in \Phi_B$ such that $A_\alpha \cap C \neq \{0\}$. Let $x_\alpha$ be a nonzero element of $A_\alpha \cap C$, and let $y_\alpha \in A_{-\alpha}$ and $z_\alpha \in B$ be as before. Observe that $z_\alpha \in C$, because of (17.15.3) and the hypothesis that $C$ be an ideal in $A$. Remember that

$$ [z_\alpha, y_\alpha]_A = -\alpha(z_\alpha) y_\alpha, $$

as in Section 17.2.
because $y_\alpha \in A_{-\alpha}$ and $z_\alpha \in B$. This implies that $y_\alpha \in C$ too, because of (17.15.4). It follows that $z_\alpha = 0$, by (17.15.3) and the commutativity of $C$ as a Lie algebra. This contradicts (17.15.4), as desired.

This shows that $A_\alpha \cap C = \{0\}$ for every $\alpha \in \Phi_B$, so that

\[(17.15.6) \quad C \subseteq A_0 = B.\]

Let $\alpha \in \Phi_B$ and $x_\alpha \in A_\alpha$ be given, so that

\[(17.15.7) \quad [w, x_\alpha]_A = \alpha(w) x_\alpha\]

for every $w \in B$. If $w \in C$, then $[w, x_\alpha]_A \in C$, because $C$ is an ideal in $A$. Combining this with (17.15.7), we get that

\[(17.15.8) \quad [w, x_\alpha]_A = 0,\]

because $A_0 \cap A_\alpha = \{0\}$ when $\alpha \neq 0$. If $x_0 \in A_0$, then $[w, x_0]_A = 0$ for every $w \in B$, and hence every $w \in C$, by (17.15.6). This implies that $C$ is contained in the center of $A$, because $A$ is spanned by $A_\alpha$, $\alpha \in \Phi_B \cup \{0\}$. It follows that $C = \{0\}$, by (17.15.2), as desired.
Chapter 18

Cartan subalgebras

18.1 Subspaces related to $\text{ad}_x$

Let $k$ be a field, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. Also let $x \in A$ be given, so that $\text{ad}_x(w) = [x, w]_A$ defines a linear mapping from $A$ into itself, as usual. If $\alpha \in k$, then put

$$A_{x, \alpha} = \{ y \in A : (\text{ad}_x - \alpha I)^l(y) = 0 \text{ for some } l \in \mathbb{Z}_+ \},$$

where $I$ is the identity mapping on $A$. This is a linear subspace of $A$, as in Section 10.7. Note that $A_{x, \alpha} \neq \{0\}$ if and only if $\alpha$ is an eigenvalue of $\text{ad}_x$ on $A$, as before.

Let $y, z \in A$ and $\alpha, \beta \in k$ be given. Remember that $\text{ad}_x$ is a derivation on $A$, as in Section 2.5. It follows that

$$\text{ad}_x - (\alpha + \beta)I)([y, z]_A) = \{\text{ad}_x(y), z\}_A + [y, \text{ad}_x(z)]_A - \alpha [y, z]_A - \beta [y, z]_A$$

as in Section 10.9. If $y \in A_{x, \alpha}$ and $z \in A_{x, \beta}$, then one can use this repeatedly to get that

$$\text{ad}_x - (\alpha + \beta)I^j([y, z]_A) = 0$$

when $j$ is sufficiently large. This means that

$$[y, z]_A \in A_{x, \alpha + \beta},$$

as in Section 10.9. In particular, $A_{x, 0}$ is a Lie subalgebra of $A$. This corresponds to the first two assertions in the lemma at the bottom of p78 of [13], and to parts (b) and (c) of Proposition 2 on p11 of [23].

Let $\alpha_1, \ldots, \alpha_n \in k$ be finitely many distinct eigenvalues of $\text{ad}_x$, and suppose that $y_j \in A_{x, \alpha_j}$ for each $j = 1, \ldots, n$. If $\sum_{j=1}^n y_j = 0$, then $y_j = 0$ for each $j = 1, \ldots, n$, as in Section 10.7.

Suppose from now on in this section that $k$ is algebraically closed, and that $A$ has positive finite dimension as a vector space over $k$. In this case, $A$ corresponds
18.2 Some remarks about characteristic polynomials

Let $k$ be a field, and let $V$ be a vector space over $k$ of positive finite dimension. Also let $a$ be a linear mapping from $V$ into itself, and let $T$ be an indeterminate. As usual, the characteristic polynomial of $a$ is defined as a formal polynomial in $T$ with coefficients in $k$ by

$$p_a(T) = \det(a - T I),$$

where $I$ is the identity mapping on $V$. More precisely, one can choose a basis for $V$, so that $a$ corresponds to a matrix with entries in $k$, and $a - T I$ corresponds to a matrix with entries in the algebra $k[T]$ of formal polynomials in $T$ with coefficients in $k$. The determinant of this matrix is an element of $k[T]$, which does not depend on the choice of basis for $V$.

Let $V_0$ be a proper nontrivial linear subspace of $V$, and suppose that

$$a(V_0) \subseteq V_0.$$  

Let $V_1$ be the quotient space $V/V_0$, so that $a$ induces a linear mapping $a_1$ from $V_1$ into itself. If $a_0$ is the restriction of $a$ to $V_0$, then

$$p_a(T) = p_{a_0}(T) p_{a_1}(T),$$

where $p_{a_0}(T), p_{a_1}(T)$ are the characteristic polynomials of $a_0, a_1$ as linear mappings on $V_0, V_1$, respectively. This can be seen using a basis for $V$ that contains a basis for $V_0$. The basis vectors not in $V_0$ can be mapped into $V_1$ by the canonical quotient mapping, to get a basis for $V_1$.

Remember that

$$V_0 = \{v \in V : a_l(v) = 0 \text{ for some } l \in \mathbb{Z}_+\}$$

is a linear subspace of $V$, as in Section 10.7. Clearly $V_0$ satisfies (18.2.2). If $v \in V$ and $a_l(v) \in V_0$, then it is easy to see that $v \in V_0$ too. This implies that the kernel of the induced mapping $a_1$ on $V_1 = V/V_0$ is trivial. Note that $V_0 \neq \{0\}$ exactly when 0 is an eigenvalue of $a$ on $V$, as in Section 10.7.

Let $n_0$ be the dimension of $V_0$, as a vector space over $k$. If $n_0 \geq 1$, then it is well known that

$$p_{a_0}(T) = (-1)^{n_0} T^{n_0}.$$
Suppose for the moment that $V_0 \neq V$, so that $V_1 \neq \{0\}$. Of course, the constant term in $p_a(T)$ is the same as the determinant of $a_1$ on $V$. This is a nonzero element of $k$, because $a_1$ is invertible on $V_1$, as in the preceding paragraph.

If $V_0 = V$, then $p_a(T) = p_m(T)$ is as in (18.2.5). If $V_0 = \{0\}$, then 0 is not an eigenvalue of $a$ on $V$, and the constant term in $p_a(T)$ is nonzero.

18.3 Characteristic polynomials and $\text{ad}_x$

Let $k$ be a field, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$, with positive finite dimension $n$ as a vector space over $k$. Also let $T$ be an indeterminate, so that the characteristic polynomial of a linear mapping from $A$ into itself is defined as a formal polynomial in $T$ with coefficients in $k$, as in the previous section. If $x \in A$, then let

$$Q_x(T) = \det(\text{ad}_x - TI)$$

be the characteristic polynomial of $\text{ad}_x$, as a linear mapping from $A$ into itself. This is the same as $(-1)^n$ times the polynomial $P_x(T)$ defined on p10 of [23].

Using a basis for $A$, we can identify $A$ with the space $k^n$ of $n$-tuples of elements of $k$, as a vector space over $k$. Of course, $\text{ad}_x$ depends linearly on $x$, and can be expressed in terms of the coordinates of $x$ and the corresponding structure constants for the Lie bracket on $A$. Similarly, (18.3.1) can be expressed as

$$Q_x(T) = \sum_{j=0}^n q_j(x) T^j,$$

where $q_j(x)$ corresponds to a homogeneous polynomial of degree $n - j$ in the coordinates of $x$ for each $j = 0, 1, \ldots, n$, as on p11 of [23]. More precisely, the coefficients for $q_j$ can be given in terms of the structure constants for the Lie bracket on $A$. By construction, $q_n(x)$ corresponds to the constant polynomial $(-1)^n$.

Suppose for the moment that $A$ is nilpotent as a Lie algebra over $k$. In this case, we can choose a basis for $A$ such that $\text{ad}_x$ always corresponds to a strictly upper-triangular matrix. This implies that the coefficients of $q_j(x)$ are all equal to 0 when $j < n$.

Let $x \in A$ be given, and let $A_{x,0}$ be as in (18.1.1). This is the same as (18.2.4), with $V = A$ and $a = \text{ad}_x$. Let $n_0(x)$ be the dimension of $A_{x,0}$. Note that $x \in A_{x,0}$, and that $A_{x,0} = A$ when $x = 0$, so that $n_0(x) \geq 1$. We also have that

$$q_j(x) = 0 \quad \text{when } j < n_0(x),$$

and

$$q_{n_0(x)}(x) \neq 0,$$

as in the previous section.

Of course, $n_0(x) = n$ exactly when $\text{ad}_x$ is nilpotent on $A$. If this happens for every $x \in A$, then $A$ is nilpotent as a Lie algebra, as in Section 9.10. These remarks correspond to some of those on p11 in [23].
18.4 Engel subalgebras

Let $k$ be a field, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. If $x \in A$, then

$$A_{x,0} = \{ y \in A : (\mathrm{ad}_x)^l(y) = 0 \text{ for some } l \in \mathbb{Z}_+ \}$$

(18.4.1)

is a Lie subalgebra of $A$, as in Section 18.1. A Lie subalgebra of $A$ of this form is called an Engel subalgebra, following Barnes, as on p79 of [13].

Let us suppose from now on in this section that $A$ has positive finite dimension $n$ as a vector space over $k$, and that the number of elements of $k$ is strictly larger than $n$, as in Exercise 5 on p81 of [13]. Let $C$ be a Lie subalgebra of $A$, and let $z$ be an element of $C$ such that $A_{z,0}$ is minimal with respect to inclusion among the $A_{x,0}$'s with $x \in C$. Of course, it suffices to choose $z \in C$ so that the dimension of $A_{z,0}$, as a vector space over $k$, is minimal. If

$$C \subseteq A_{z,0},$$

(18.4.2)

then

$$A_{z,0} \subseteq A_{z,0}$$

(18.4.3)

for every $x \in C$. This is Lemma A on p79 of [13].

We may as well suppose that $A_{z,0} \neq \{0\}$, since otherwise (18.4.3) is trivial. We may as well suppose too that $A_{z,0} \neq A$, because (18.4.3) would follow from the minimality of $A_{z,0}$ otherwise. Let $r$ be the dimension of $A_{z,0}$ as a vector space over $k$, so that $1 \leq r < n$.

Let $x \in C$ and $c \in k$ be given, and observe that

$$z + cx \in A_{z,0},$$

(18.4.4)

Thus $z + cx \in A_{z,0}$, by (18.4.2). This implies that $\mathrm{ad}_{z+cx}$ maps $A_{z,0}$ into itself, because $A_{z,0}$ is a Lie subalgebra of $A$. The quotient $A/A_{z,0}$ is defined as a vector space over $k$, and $\mathrm{ad}_{z+cx}$ induces a mapping from $A/A_{z,0}$ into itself in the usual way.

Note that $A_{z,0}$ corresponds to (18.2.4) with $V = A$ and $a = \mathrm{ad}_z$. Thus the mapping from $A/A_{z,0}$ into itself induced by $\mathrm{ad}_z$ is injective, as before.

Let $T$ be an indeterminate, and let $h(T, c)$ be the characteristic polynomial of $\mathrm{ad}_{z+cx}$ on $A$. As in Section 18.2, this can be expressed as

$$h(T, c) = f(T, c)g(T, c),$$

(18.4.5)

where $f(T, c)$ is the characteristic polynomial of the restriction of $\mathrm{ad}_{z+cx}$ to $A_{z,0}$, and $g(T, c)$ is the characteristic polynomial of the linear mapping from $A/A_{z,0}$ into itself induced by $\mathrm{ad}_{z+cx}$. These polynomials can be expressed as

$$f(T, c) = (-1)^r T^r + f_1(c) T^{r-1} + \cdots + f_r(c)$$

(18.4.6)

and

$$g(T, c) = (-1)^{n-r} T^{n-r} + g_1(c) T^{n-r-1} + \cdots + g_{n-r}(c),$$

(18.4.7)
CHAPTER 18. CARTAN SUBALGEBRAS

as on p79 of [13]. More precisely, \( f_j(c) \) and \( g_j(c) \) can be given by polynomials with coefficients in \( k \), of degree less than or equal to \( j \).

Observe that \( g_{n-r}(0) \neq 0 \), because the mapping induced on \( A/A_{z,0} \) by \( \text{ad}_z \) is invertible. In particular, \( g_{n-r} \neq 0 \) on \( k \), so that \( g_{n-r} \) has at most \( n - r \) zeros in \( k \). Remember that \( k \) has at least \( n + 1 \) elements, by hypothesis. It follows that there are \( r + 1 \) distinct elements \( c_1, \ldots, c_{r+1} \) of \( k \) such that

\[
g_{n-r}(c_l) \neq 0
\]

for each \( l = 1, \ldots, r + 1 \).

By construction, \( g_{n-r}(c) \) is the determinant of the linear mapping from \( A/A_{z,0} \) into itself induced by \( \text{ad}_{z+cx} \). Suppose for the moment that \( c \in k \) has the property that \( g_{n-r}(c) \neq 0 \), so that the mapping from \( A/A_{z,0} \) into itself induced by \( \text{ad}_{z+cx} \) is injective. Of course, this implies that all positive powers of this mapping are injective. Remember that the positive powers of the mapping from \( A/A_{z,0} \) induced by \( \text{ad}_{z+cx} \) correspond to the mappings from \( A/A_{z,0} \) into itself induced by the positive powers of \( \text{ad}_{z+cx} \). It follows that the kernels of the positive powers of \( \text{ad}_{z+cx} \) on \( A \) are contained \( A_{z,0} \). This implies that

\[
A_{z+cx,0} \subseteq A_{z,0}
\]

Under these conditions, we get that

\[
A_{z+cx,0} = A_{z,0}
\]

because of the minimality condition on \( A_{z,0} \).

As before, \( A_{z+cx,0} \) corresponds to (18.2.4) with \( V = A \) and \( a = \text{ad}_{z+cx} \). If (18.4.10) holds, then the characteristic polynomial \( f(T,c) \) of the restriction of \( \text{ad}_{z+cx} \) to (18.4.10) is equal to \((-1)^r T^r \). This means that

\[
f_j(c) = 0
\]

for each \( j = 1, \ldots, r \). In particular, this holds for \( r + 1 \) distinct elements of \( k \), as in (18.4.8). This implies that

\[
f_j \equiv 0
\]

on \( k \) for each \( j = 1, \ldots, r \), because \( f_j \) is given by a polynomial of degree less than or equal to \( j \).

Equivalently,

\[
f(T, c) = (-1)^r T^r
\]

for every \( c \in k \). Let \( c \in k \) be given again, and consider

\[
A_{z+cx,0} \cap A_{z,0}
\]

which is a linear subspace of \( A_{z,0} \). If \( V = A_{z,0} \) and \( a = \text{ad}_{z+cx} \), then (18.4.14) corresponds to (18.2.4). Using (18.4.13), we get that (18.4.14) is the same as \( A_{z,0} \), as in Section 18.2. This means that

\[
A_{z,0} \subseteq A_{z+cx,0}
\]
for every $c \in k$.

Remember that this holds for every $x \in C$. Thus we can replace $x$ with $x - z$
and take $c = 1$ to get (18.4.3), as desired.

18.5 Self-normalizing subalgebras

Let $k$ be a field again, and let $(A, [\cdot, \cdot])$ be a Lie algebra over $k$. Also let
$C$ be a Lie subalgebra of $A$, and remember that the normalizer $N_A(C)$ of $C$ in $A$ is
the set of $w \in A$ such that $[w, z]_A \in C$ for every $z \in C$. This is the largest Lie
subalgebra of $A$ that contains $C$ as an ideal, as in Section 9.8. If $N_A(C) = C$,
then $C$ is said to be self-normalizing in $A$, as before.

Let $x \in A$ be given, and let $A_{x,0}$ be the set of $y \in A$ such that
$(\text{ad}_x)^l(y) = 0$ for some positive integer $l$, as before. This is a Lie subalgebra of $A$, as in Section
18.1. Equivalently, if $V = A$ and $a = \text{ad}_x$, then $A_{x,0}$ is the same as (18.2.4).
Remember that the mapping from $A/A_{x,0}$ into itself induced by $\text{ad}_x$ is injective,
as in Section 18.2.

Let $w$ be an element of the normalizer $N_A(A_{x,0})$ of $A_{x,0}$ in $A$. Note that
$x \in A_{x,0}$, so that
$$[w, x]_A \in A_{x,0}. \quad (18.5.1)$$
This means that $(\text{ad}_x)^l([w, x]_A) = 0$ for some positive integer $l$, and hence
$(\text{ad}_x)^{l+1}(w) = 0$. It follows that $w \in A_{x,0}$, so that
$$N_A(A_{x,0}) = A_{x,0}. \quad (18.5.2)$$
This corresponds to the last part of the proof of Theorem 1 on p12 of [23].

Suppose from now on in this section that the dimension of $A$ is finite, as a
vector space over $k$. Thus the dimension of the quotient space $A/A_{x,0}$ is finite
too. It follows that the mapping on $A/A_{x,0}$ induced by $\text{ad}_x$ is injective, because
it is injective, as before.

Suppose that $C$ is a Lie subalgebra of $A$ with
$$A_{x,0} \subseteq C. \quad (18.5.3)$$
In particular, $x \in C$, because $x \in A_{x,0}$. Thus $\text{ad}_x$ maps $C$ into itself, so that
the mapping on $A/A_{x,0}$ induced by $\text{ad}_x$ maps $C/A_{x,0}$ into itself. More precisely,
the mapping induced on $A/A_{x,0}$ maps $C/A_{x,0}$ onto itself, because this mapping
is injective, and the dimension of $C/A_{x,0}$ is finite.

The normalizer $N_A(C)$ of $C$ in $A$ is mapped into $C$ by $\text{ad}_x$, because $x \in C$.
This implies that the mapping induced on $A/A_{x,0}$ by $\text{ad}_x$ maps $N_A(C)/A_{x,0}$
to $C/A_{x,0}$. It follows that
$$N_A(C)/A_{x,0} = C/A_{x,0}, \quad (18.5.4)$$
because the mapping induced on $A/A_{x,0}$ by $\text{ad}_x$ is injective and maps $C/A_{x,0}$
onto itself.
This means that
\[ N_A(C) = C, \]
because of (18.5.3) and the fact that \( C \subseteq N_A(C) \). This corresponds to Lemma B on p79 of [13].

18.6 Orthogonality conditions

Let \( k \) be a field, and let \( (A, [\cdot, \cdot]) \) be a Lie algebra over \( k \). Also let \( b(\cdot, \cdot) \) be a bilinear form on \( A \) that is associative, or equivalently invariant under the adjoint representation. This means that
\[ b([w, y]_A, z) = -b(y, [w, z]_A) \]
for every \( w, y, z \in A \), as in Sections 6.10 and 7.7. Let \( x \in A \) be given, so that
\[ b(\text{ad}_x(y), z) + b(y, \text{ad}_x(z)) = 0 \]
for every \( y, z \in A \). If \( \alpha, \beta \in k \), then we get that
\[ b((\text{ad}_x - \alpha I)(y), z) + b(y, (\text{ad}_x - \beta I)(z)) = -(\alpha + \beta) b(y, z) \]
for every \( y, z \in A \).

Let \( l \) be a positive integer, and consider
\[ b((\text{ad}_x - I)^l(y), (\text{ad}_x - \beta I)^l(z)). \]
If \( l_1 \) and \( l_2 \) are nonnegative integers, then consider
\[ b((\text{ad}_x - \alpha I)^{l_1}(y), (\text{ad}_x - \beta I)^{l_2}(z)). \]
If \( l_1 \) or \( l_2 \) is equal to 0, then \((\text{ad}_x - \alpha I)^{l_1} \) or \((\text{ad}_x - \beta I)^{l_2} \) is interpreted as being the identity operator on \( A \), as appropriate. We can express (18.6.4) as a sum of terms of the form (18.6.5), using (18.6.3) repeatedly. More precisely, we have that
\[ l_1 + l_2 = l \]
in each of the terms of the form (18.6.5).

Remember that \( A_{x, \alpha} \) consists of the \( y \in A \) such that \((\text{ad}_x - \alpha I)^y(y) = 0 \) for all sufficiently large positive integers \( j \), as in Section 18.1, and similarly for \( \beta \). If \( y \in A_{x, \alpha} \) and \( z \in A_{x, \beta} \), then (18.6.5) is equal to 0 when \( l_1 \) or \( l_2 \) is sufficiently large. This implies that
\[ (\alpha + \beta)^l b(y, z) = 0 \]
when \( l \) is sufficiently large. It follows that
\[ b(y, z) = 0 \]
when \( \alpha + \beta \neq 0 \). This corresponds to a remark in the proof of part (d) of Theorem 3 on p15 of [23].
Remember that $A_{x, \alpha} \neq \{0\}$ exactly when $\alpha \in k$ is an eigenvalue of $\text{ad}_x$ on $A$. Suppose now that $k$ is algebraically closed, and that $A$ has positive finite dimension as a vector space over $k$. In this case, $A$ corresponds to the direct sum of the nonzero subspaces $A_{x, \alpha}$, as a vector space over $k$.

Suppose that $b(\cdot, \cdot)$ is also nondegenerate on $A$. Under these conditions, it is easy to see that the restriction of $b(\cdot, \cdot)$ to $A_{x, \alpha}$ is nondegenerate, using (18.6.8). This corresponds to part (d) of Theorem 3 on p15 of [23].

18.7 NILPOTENCE AND NORMALIZING ELEMENTS

Let $k$ be a field, and let $(A, [\cdot, \cdot]_A)$ be a finite-dimensional Lie algebra over $k$. Also let $C_1, C_2$ be Lie subalgebras of $A$, with

$$C_1 \subseteq C_2.$$  

Thus the quotient $C_2/C_1$ is defined as a vector space over $k$, and we let $q$ be the canonical quotient mapping from $C_2$ onto $C_2/C_1$.

If $x \in C_1$, then $\text{ad}_x$ maps $C_1$ and $C_2$ into themselves. This leads to a linear mapping $\text{ad}_{C_2/C_1}^x$ from $C_2/C_1$ into itself, with

$$\text{ad}_{C_2/C_1}^x \circ q = q \circ \text{ad}_x.$$  

It is easy to see that $x \mapsto \text{ad}_{C_2/C_1}^x$ defines a representation of $C_1$, as a Lie algebra over $k$, on $C_2/C_1$. In particular,

$$\{\text{ad}_{C_2/C_1}^x : x \in C_1\}$$

is a Lie subalgebra of the Lie algebra $gl(C_2/C_1)$ of linear mappings from $C_2/C_1$ into itself, with respect to the usual commutator bracket.

Observe that

$$\text{ad}_{C_2/C_1}^x \circ q = 0.$$

for some positive integer $l$ if and only if

$$q \circ (\text{ad}_x)^l = 0.$$  

Equivalently, this means that

$$\text{ad}_x^l(C_2) \subseteq C_1.$$  

In this case, if $\text{ad}_x$ is nilpotent on $C_1$, then it follows that $\text{ad}_x$ is nilpotent on $C_2$. Of course, if $\text{ad}_x$ is nilpotent on $C_2$, then $\text{ad}_{C_2/C_1}^x$ is nilpotent on $C_2/C_1$.

Suppose that $C_1 \neq C_2$, so that $C_2/C_1 \neq \{0\}$, and that $\text{ad}_{C_2/C_1}^x$ is nilpotent on $C_2/C_1$ for each $x \in C_1$. Under these conditions, there is a nonzero element of $C_2/C_1$ that is mapped to 0 by $\text{ad}_{C_2/C_1}^x$ for every $x \in C_1$, as in Section 9.9. This means that there is a $y \in C_2$ such that $y \notin C_1$ and

$$\text{ad}_{C_2/C_1}^x(q(y)) = 0.$$  

for every $x \in X_1$, and hence
\[(18.7.8) \quad q(ad_x(y)) = 0\]
for every $x \in C_1$. This is the same as saying that
\[(18.7.9) \quad ad_x(y) = [x, y]_A \in C_1\]
for every $x \in C_1$. Equivalently, this means that $y$ is an element of the normalizer $N_A(C_1)$ of $C_1$ in $A$, as in Section 9.8.

If $C_2$ is nilpotent as a Lie algebra over $k$, then $ad_x$ is nilpotent on $C_2$ for every $x \in C_2$. In particular, this implies that $ad_x^{C_2/C_1}$ is nilpotent on $C_2/C_1$ for every $x \in C_1$, as before. If $C_1$ is a proper Lie subalgebra of $C_2$, then it follows that $C_1$ is a proper subset of its normalizer $N_{C_2}(C_1)$ in $C_2$, as in the preceding paragraph. This corresponds to Exercise 7 on p14 of [13].

### 18.8 Minimal Engel subalgebras

Let $k$ be a field, and let $(A, [, ]_A)$ be a Lie algebra over $k$. A Lie subalgebra $C$ of $A$ is said to be a Cartan subalgebra if $C$ is nilpotent as a Lie algebra over $k$, and $C$ is self-normalizing in $A$, as on p80 of [13], and Definition 1 on p10 of [23]. Remember that the normalizer $N_A(C)$ of $C$ in $A$ is the largest Lie subalgebra of $A$ that contains $C$ as an ideal, and that $C$ is self-normalizing in $A$ when $N_A(C) = C$, as in Section 9.8.

Let us suppose from now on in this section that $A$ has finite dimension as a vector space over $k$, and that the number of elements of $k$ is strictly larger than the dimension of $A$ as a vector space over $k$. If $x \in A$, then $A_{x,0}$ is the set of $y \in A$ such that $(ad_x)^l(y) = 0$ for some positive integer $l$, as in Section 18.1. This is a Lie subalgebra of $A$, which is known as an Engel subalgebra of $A$, as before. Remember that Engel subalgebras of $A$ are self-normalizing in $A$, as in Section 18.5.

Suppose that $z \in A$ has the property that $A_{z,0}$ is minimal with respect to inclusion among Engel subalgebras of $A$. In particular, one could take $A_{z,0}$ to be an Engel subalgebra of minimal dimension, as a vector space over $k$. Under these conditions, $A_{z,0}$ is a Cartan subalgebra of $A$, as in the theorem on p80 of [13]. This is related to Theorem 1 on p12 of [23] when $k = C$.

It suffices to show that $A_{z,0}$ is nilpotent as a Lie algebra over $k$, because $A_{z,0}$ is self-normalizing in $A$, as before. Note that $C = A_{z,0}$ satisfies the conditions in Section 18.4, because $A_{z,0}$ is minimal, by hypothesis. If $x \in A_{z,0}$, then we get that
\[(18.8.1) \quad A_{z,0} \subseteq A_{x,0},\]
as before. Thus, for each $y \in A_{z,0}$, we have that $(ad_x)^l(y) = 0$ for some positive integer $l$. This implies that the restriction of $ad_x$ to $A_{z,0}$ is nilpotent as a linear mapping from $A_{z,0}$ into itself. Equivalently, this means that $x$ is ad-nilpotent as an element of $A_{z,0}$, considered as a Lie algebra over $k$. It follows that $A_{z,0}$ is nilpotent as a Lie algebra over $k$, as in Section 9.10.
Let $C$ be a Lie subalgebra of $A$. If $x \in C$, then $\text{ad}_{C,x}$ is the same as the restriction of $\text{ad}_x$ to $C$. Suppose that $C$ is nilpotent as a Lie algebra over $k$, so that $\text{ad}_{C,x}$ is nilpotent on $C$ for every $x \in C$. This implies that
\begin{equation}
C \subseteq A_{x,0}
\end{equation}
for every $x \in C$. In particular, if $C$ is an Engel subalgebra of $A$, then it follows that $C$ is minimal among Engel subalgebras of $A$.

If $C$ is a Cartan subalgebra of $A$, then the other part of the theorem on p80 of [13] states that $C$ is a minimal Engel subalgebra of $A$. This corresponds to Corollary 2 on p13 of [23] when $k = C$. Of course, (18.8.2) holds for every $x \in C$ in this case, because $C$ is nilpotent. Suppose for the sake of a contradiction that for each $x \in C$, $C \neq A_{x,0}$. Let $z$ be an element of $C$ be such that $A_{z,0}$ is minimal with respect to inclusion. As usual, one can take $z \in C$ such that $A_{z,0}$ has minimal dimension, as a vector space over $k$. Note that
\begin{equation}
C \subseteq A_{z,0},
\end{equation}
by (18.8.2). Under these conditions, we get that (18.8.1) holds for every $x \in C$, as in Section 18.4 again.

Let $x \in C$ be given. Of course, $\text{ad}_x$ maps $C$ into itself, because $C$ is a Lie subalgebra of $A$. Observe that $x \in A_{z,0}$, by (18.8.3). Thus $\text{ad}_x$ maps $A_{z,0}$ into itself, because $A_{z,0}$ is a Lie subalgebra of $A$ too. The restriction of $\text{ad}_x$ to $A_{z,0}$ is nilpotent, because of (18.8.1).

Consider the quotient $A_{z,0}/C$, as a finite-dimensional vector space over $k$. If $x \in C$, then $\text{ad}_x$ induces a linear mapping from $A_{z,0}/C$ into itself, by the remarks in the preceding paragraph. This defines a representation of $C$, as a Lie algebra over $k$, on $A_{z,0}/C$. The mapping induced on $A_{z,0}/C$ by $\text{ad}_x$ is nilpotent for each $x \in C$, because $\text{ad}_x$ is nilpotent on $A_{z,0}$.

By hypothesis, $C \neq A_{z,0}$, so that $A_{z,0}/C \neq \{0\}$. It follows that there is a nonzero element of $A_{z,0}/C$ that is mapped to 0 by the mapping induced by $\text{ad}_x$ for every $x \in C$, as in Section 9.9. Equivalently, this means that there is a $y \in A_{z,0}$ such that $y \notin C$ and
\begin{equation}
\text{ad}_x(y) = [x,y]_A \in C
\end{equation}
for every $x \in C$. This means that $y \in N_A(C)$, which is a contradiction, because $C$ is supposed to be self-normalizing in $A$. This is another instance of the argument in the previous section.

18.9 Nilpotent vectors and sums

Let $k$ be a field, and let $V$ be a vector space over $k$. If $T$ is a linear mapping from $V$ into itself, then let $E_0(T)$ be the set of $v \in V$ such that $T^l(v) = 0$ for some positive integer $l$, as in Section 10.7. Remember that this is a linear subspace of
V, as before. Let R be another linear mapping from V into itself, and suppose that R commutes with T. Under these conditions, one can check that
\[\mathcal{E}_0(R) \cap \mathcal{E}_0(T) \subseteq \mathcal{E}_0(R + T).\]

More precisely, if \(v \in V\) and \(l \in \mathbb{Z}_+\), then
\[\mathcal{E}_0(R + T)^l(v)\]
can be expressed as a sum of terms of the form
\[R^l(T^l(v)) = T^l(R^l(v)),\]
where \(l_1, l_2\) are nonnegative integers such that \(l_1 + l_2 = l\). If \(v\) is an element of the left side of (18.9.1), then (18.9.3) is equal to 0 when \(l_1\) or \(l_2\) is sufficiently large. This implies that (18.9.2) is equal to 0 when \(l\) is sufficiently large, as desired.

Suppose for the moment that \(\mathcal{E}_0(R) = V\);
\[\mathcal{E}_0(T) \subseteq \mathcal{E}_0(R + T),\]
by (18.9.1). Similarly,
\[\mathcal{E}_0(R + T) \subseteq \mathcal{E}_0(T),\]
by considering T as \((R + T) + (-R)\). Thus
\[\mathcal{E}_0(T) = \mathcal{E}_0(T + R)\]
under these conditions. If T is diagonalizable on V, then one can check that \(\mathcal{E}_0(T)\) is the same as the kernel of T.

Let \((A, [,]_A)\) be a Lie algebra over \(k\). If \(x \in A\), then
\[A_{x,0} = \mathcal{E}_0(\text{ad}_x)\]
in the notation of Section 18.1, with \(V = A\) as a vector space over \(k\) on the right side. Suppose that \(y \in A\) satisfies \([x, y]_A = 0\), so that \(\text{ad}_x\) and \(\text{ad}_y\) commute on \(A\), as in Section 2.4. This implies that
\[A_{x,0} \cap A_{y,0} \subseteq A_{x+y,0},\]
as in (18.9.1). If \(A_{y,0} = A\), then we get that
\[A_{x,0} = A_{x+y,0},\]
as in (18.9.7). If \(\text{ad}_x\) is diagonalizable on \(A\), then \(A_{x,0}\) is the same as the kernel of \(\text{ad}_x\) on \(A\), as before. This means that
\[A_{x,0} = C_A(x) = C_A(\{x\})\]
18.10 CARTAN SUBALGEBRAS AND SEMISIMPLICITY

is the centralizer of \( x \) in \( A \), as in Section 7.6.

Suppose that \( k \) is an algebraically closed field of characteristic 0, and that \( A \) is a finite-dimensional semisimple Lie algebra over \( k \). Let \( x \in A \) be given, and let \( x = x_1 + x_2 \) be the abstract Jordan decomposition of \( x \) in \( A \), as in Section 14.3. Thus \( x_1 \in A \) is ad-diagonalizable, \( x_2 \in A \) is ad-nilpotent, and \([x_1, x_2]_A = 0\). It follows that

\[
A_{x,0} = A_{x_1,0} = C_A(x_1),
\]

as in (18.9.10) and (18.9.11). This corresponds to some of the arguments in the proof of the corollary on p80 in [13].

18.10 Cartan subalgebras and semisimplicity

Let \( k \) be an algebraically closed field of characteristic 0, and let \((A, [\cdot,\cdot]_A)\) be a finite-dimensional semisimple Lie algebra over \( k \). If \( B \) is a maximal toral subalgebra of \( A \), then \( B \) is commutative as a Lie algebra, and the centralizer \( C_A(B) \) of \( B \) in \( A \) is equal to \( B \), as in Sections 17.1 and 17.4. The normalizer \( N_A(B) \) of \( B \) in \( A \) is equal to \( B \) too, as in Section 17.14. It follows that \( B \) is a Cartan subalgebra of \( A \), as mentioned at the top of p80 of [13].

Conversely, let \( C \) be any Cartan subalgebra of \( A \). In particular, \( C \) is an Engel subalgebra of \( A \), as in Section 18.8. Thus there is an \( x \in A \) such that

\[
C = A_{x,0},
\]

and one can take \( x \) to be ad-diagonalizable, as in the previous section. This means that the linear span of \( x \) in \( A \) is a toral subalgebra of \( A \). Let \( B \) be a maximal toral subalgebra of \( A \) that contains \( x \), which can be obtained by taking \( B \) to have maximal dimension. Observe that

\[
B \subseteq C_A(x) = C,
\]

because \( B \) is commutative as a Lie algebra, as in Section 17.1. However, \( B \) is a Cartan subalgebra of \( A \), as in the preceding paragraph, and hence an Engel subalgebra of \( A \), as in Section 18.8. It follows that

\[
B = C,
\]

because \( C \) is a minimal Engel subalgebra of \( A \), as in Section 18.8. This is the corollary on p80 of [13].

Alternatively, suppose for the moment that \( k \) is any field of characteristic 0, and that \( A \) is a finite-dimensional semisimple Lie algebra over \( k \). Suppose also that (18.10.1) is a Cartan subalgebra of \( A \) for some \( x \in A \). Remember that the Killing form on \( A \) is nondegenerate, as in Section 10.13. This implies that the restriction of the Killing form on \( A \) to \( C \) is nondegenerate, as in Section 18.6, and part (d) of Theorem 3 on p15 of [23].
CHAPTER 18. CARTAN SUBALGEBRAS

Remember that the collection of linear mappings on $A$ of the form $\text{ad}_w$, $w \in A$, is a Lie subalgebra of the Lie algebra $gl(A)$ of all linear mappings from $A$ into itself, with respect to the usual commutator bracket. This is the same as the image of $A$ under the adjoint representation. Similarly,

\begin{equation}
\{\text{ad}_w : w \in C\}
\end{equation}

is a Lie subalgebra of the Lie algebra just mentioned, because $C$ is a Lie subalgebra of $A$. More precisely, $(18.10.4)$ is nilpotent as a Lie algebra, because $C$ is nilpotent, by hypothesis. In particular, $(18.10.4)$ is solvable as a Lie algebra over $k$.

If $w, y, z \in C$, then it follows that

\begin{equation}
\text{tr}_A(\text{ad}_w \circ \text{ad}_{[y,z]}) = 0,
\end{equation}

as in Section 10.5. This uses the fact that $\text{ad}_{[y,z]} = [\text{ad}_y, \text{ad}_z]$, as in Section 2.4. This implies that

\begin{equation}
[y, z] = 0
\end{equation}

for every $y, z \in C$, because the restriction of the Killing form on $A$ to $C$ is nondegenerate, as before. This shows that $C$ is commutative as a Lie algebra over $k$, as in part (a) of Theorem 3 on p15 of [23].

It follows that $C$ is contained in its centralizer in $A$. The centralizer of $C$ in $A$ is automatically contained in the normalizer $N_A(C)$ of $C$ in $A$, as in Section 17.14. In this situation, $N_A(C) = C$, so that $C$ is equal to its centralizer in $A$. This is part (b) of Theorem 3 on p80 of [23]. This implies that $C$ is maximal among commutative Lie subalgebras of $A$, as in Corollary 1 on p15 of [23].

Suppose now that $k$ is also algebraically closed again. Let $w \in C$ be given, and let

\begin{equation}
w = w_1 + w_2
\end{equation}

be the abstract Jordan decomposition of $w$ in $A$, as in Section 14.3. This means that $w_1 \in A$ is ad-diagonalizable, $w_2 \in A$ is ad-nilpotent, and $[w_1, w_2] = 0$. Let $y \in C$ be given, so that $[w, y] = 0$, and hence $\text{ad}_w$ commutes with $\text{ad}_y$ on $A$, as in Section 2.4. This implies that $\text{ad}_y$ commutes with $\text{ad}_w$, as in Section 10.8. It follows that

\begin{equation}
[w_1, y] = [w_2, y] = 0.
\end{equation}

More precisely, this uses the fact that the kernel of the adjoint representation of $A$ is the center of $A$ as a Lie algebra, which is trivial in this case, because $A$ is semisimple by hypothesis.

Thus $\text{ad}_{w_2}$ commutes with $\text{ad}_y$ on $A$, as in Section 2.4. This implies that $\text{ad}_{w_2} \circ \text{ad}_y$ is nilpotent on $A$, because $\text{ad}_{w_2}$ is nilpotent on $A$. It follows that

\begin{equation}
\text{tr}_A(\text{ad}_{w_2} \circ \text{ad}_y) = 0.
\end{equation}

This shows that $w_2 = 0$, because the restriction of the Killing form on $A$ to $C$ is nondegenerate, as before. This means that $w = w_1$ is ad-diagonalizable as an element of $A$, as in part (c) of Theorem 3 on p15 of [23].
The hypothesis that (18.10.1) be a Cartan subalgebra of \( A \) really means that \( C \) is nilpotent as a Lie algebra over \( k \), because Engel subalgebras of \( A \) are self-normalizing, as in Section 18.5. In the previous arguments, it is enough to ask that \( C \) be solvable as a Lie algebra over \( k \), so that (18.10.4) is solvable as a Lie algebra. The same conclusion is mentioned in Exercise 2 on p81 of [13].

18.11 Subalgebras and homomorphisms

Let \( k \) be a field, and let \((\mathcal{A};\mathcal{B})\) be a Lie algebra over \( k \). If \( C \) is a Cartan subalgebra of \( A \), then \( C \) is maximal among nilpotent Lie subalgebras of \( A \). This is the first part of Exercise 4 on p81 in [13], and it follows from the statement mentioned at the end of Section 18.7.

If \( C_1, C_2 \) are Cartan subalgebras of \( A \) with

\[
(18.11.1) \quad C_1 \subseteq C_2,
\]

then

\[
(18.11.2) \quad C_1 = C_2,
\]

because \( C_1 \) is maximal among nilpotent Lie subalgebras of \( A \), as in the preceding paragraph. Alternatively, if \( C_1, C_2 \) are Engel subalgebras of \( A \), and \( C_2 \) is minimal among Engel subalgebras of \( A \), then (18.11.1) implies (18.11.2).

Let \( \tilde{A} \) and \( C \) be Lie subalgebras of \( A \), with

\[
(18.11.3) \quad C \subseteq \tilde{A}.
\]

It is easy to see that

\[
(18.11.4) \quad N_{\tilde{A}}(C) = N_A(C) \cap \tilde{A},
\]

where \( N_A(C) \), \( N_{\tilde{A}}(C) \) are the normalizers of \( C \) in \( A \) and \( \tilde{A} \), respectively. In particular, if \( C \) is self-normalizing in \( A \), then \( C \) is self-normalizing in \( \tilde{A} \). If \( C \) is a Cartan subalgebra of \( A \), then it follows that \( C \) is a Cartan subalgebra of \( \tilde{A} \).

Let \((\mathcal{A}',\mathcal{B}')\) be another Lie algebra over \( k \), and let \( \phi \) be a Lie algebra homomorphism from \( A \) onto \( A' \). If \( x \in A \) and \( x' \in A' \), then we put

\[
(18.11.5) \quad \text{ad}_{\mathcal{A},x}(y) = [x, y]_A, \quad \text{ad}_{\mathcal{A}',x'}(y') = [x', y']_{\mathcal{A}'}
\]

for every \( y \in A, y' \in A' \), as usual. Thus

\[
(18.11.6) \quad \phi(\text{ad}_{\mathcal{A},x}(y)) = \phi([x, y]_A) = [\phi(x), \phi(y)]_{\mathcal{A}'} = \text{ad}_{\mathcal{A}',\phi(x)}(\phi(y))
\]

for every \( x, y \in A \).

Let \( C' \) be a Lie subalgebra of \( A' \). If \( x \in A \), then it is easy to see that

\[
(18.11.7) \quad \text{ad}_{\mathcal{A}',\phi(x)}(C') \subseteq C'
\]

if and only if

\[
(18.11.8) \quad \text{ad}_{\mathcal{A},x}(\phi^{-1}(C')) \subseteq \phi^{-1}(C'),
\]
CHAPTER 18. CARTAN SUBALGEBRAS

using (18.11.6). This means that \( \phi(x) \) is an element of the normalizer \( N_{A'}(C') \) of \( C' \) in \( A' \) if and only if \( x \) is an element of the normalizer \( N_A(\phi^{-1}(C')) \) of \( \phi^{-1}(C') \) in \( A \). It follows that \( C' \) is self-normalizing in \( A' \) if and only if \( \phi^{-1}(C') \) is self-normalizing in \( A \).

Suppose from now on in this section that \( A \) has finite dimension as a vector space over \( k \), so that the dimension of \( A' \) is finite as well. If \( \phi^{-1}(C') \) contains an Engel subalgebra of \( A \), then \( \phi^{-1}(C') \) is self-normalizing in \( A' \), as in Section 18.5.

Let \( C \) be a Lie subalgebra of \( A \) again, so that \( \phi(C) \) is a Lie subalgebra of \( A' \). Note that \( \phi^{-1}(\phi(C)) \) is a Lie subalgebra of \( A \), which is spanned by \( C \) and the kernel of \( \phi \). If \( C \) contains an Engel subalgebra of \( A \), then \( \phi^{-1}(\phi(C)) \) contains the same Engel subalgebra of \( A \).

Suppose for the rest of the section that the number of elements of \( k \) is strictly larger than the dimension of \( A \) as a vector space over \( k \). If \( C \) is a Cartan subalgebra of \( A \), then \( C \) is an Engel subalgebra of \( A \), as in Section 18.8. It follows that \( \phi(C) \) is self-normalizing in \( A' \), as in the previous paragraphs. We also have that \( C \) is nilpotent as a Lie algebra over \( k \), so that \( \phi(C) \) is nilpotent too, as in Section 9.5. This shows that \( \phi(C) \) is a Cartan subalgebra of \( A' \), as in Lemma A on p81 of [13].

Suppose now that \( C' \) is a Cartan subalgebra of \( A' \), and that \( C \) is a Cartan subalgebra of \( \phi^{-1}(C') \), as a Lie algebra over \( k \). We would like to show that \( C \) is a Cartan subalgebra of \( A \) too, as in Lemma B on p81 of [13]. Observe that \( \phi(C) \) is a Cartan subalgebra of \( \phi(\phi^{-1}(C')) = C' \), as in the preceding paragraph. We also have that \( C' \) is a Cartan subalgebra of itself, because \( C' \) is nilpotent as a Lie algebra. This implies that

\[
(18.11.9) \quad \phi(C) = C',
\]

as in (18.11.2).

Suppose that \( x \) is an element of the normalizer \( N_A(C) \) of \( C \) in \( A \), and let us check that \( x \in C \). It is easy to see that \( \phi(x) \) is in the normalizer \( N_{A'}(\phi(C)) \) of \( \phi(C) \) in \( A' \). This implies that \( \phi(x) \in \phi(C) \), because of (18.11.9) and the hypothesis that \( C' \) be a Cartan subalgebra of \( A' \). Thus

\[
(18.11.10) \quad x \in \phi^{-1}(\phi(C)) = \phi^{-1}(C').
\]

It follows that \( x \in C \), as desired, because \( C \) is a Cartan subalgebra of \( \phi^{-1}(C') \), by hypothesis.

18.12 Unions of hyperplanes

Let \( k \) be a field, and let \( V \) be a nonzero vector space over \( k \). By a hyperplane in \( V \) we mean a linear subspace \( W \) of codimension 1 in \( V \), so that \( V \) is spanned by \( W \) and a single vector in \( V \) that is not in \( W \). Equivalently, this means that \( W \) is the kernel of a nonzero linear functional on \( V \). If \( W_1, \ldots, W_r \) are finitely many hyperplanes in \( V \), and if \( r \) is less than or equal to the number of elements
of \( k \), then we would like to check that \( \bigcup_{j=1}^{r} W_j \) is a proper subset of \( V \). Note that \( \bigcap_{j=1}^{r} W_j \) is a linear subspace of \( V \) with codimension less than or equal to \( r \). Thus the quotient space \( V/\bigcap_{j=1}^{r} W_j \) is a vector space over \( k \) of dimension less than or equal to \( r \). We may as well suppose that the dimension of \( V \) is finite, since otherwise we could replace \( V \) with this quotient.

Of course, \( k^2 = k \times k \) is a two-dimensional vector space over \( k \), with respect to coordinatewise addition and scalar multiplication. The lines in \( k^2 \) are characterized by their slope in \( k \), except for the second coordinate axis. The number of distinct lines in \( k^2 \) is the number of elements of \( k \) plus 1, which is to say that there are infinitely many distinct lines in \( k^2 \) when \( k \) has infinitely many elements. These lines are pairwise disjoint away from 0. It follows that \( k^2 \) cannot be expressed as the union of finitely many lines, where the number of lines is less than or equal to the number of elements of \( k \).

If \( V \) has dimension one, then \( \{0\} \) is the only hyperplane in \( V \), and the statement under consideration is trivial. If \( V \) has dimension at least two, then one can find at least as many distinct hyperplanes in \( V \) as lines in \( k^2 \). In particular, there is a hyperplane \( Z \) in \( V \) that is different from \( W_j \) for each \( j = 1, \ldots, r \). Thus \( Z \cap W_j \) is a hyperplane in \( Z \) for each \( j = 1, \ldots, r \). Note that the dimension of \( Z \) is one less than the dimension of \( V \), which is strictly less than the dimension of \( V \), because the dimension of \( V \) is finite. This permits us to use induction to get that the union of \( Z \cap W_j \), \( 1 \leq j \leq r \), is a proper subset of \( Z \). This implies that the union of \( W_j \), \( 1 \leq j \leq r \), is a proper subset of \( V \), as desired.

If \( a \) is a nonzero element of \( V \), then

\[
V'_a = \{ \lambda \in V' : \lambda(a) = 0 \}
\]

is a hyperplane in the dual space \( V' \) of linear functionals on \( V \). If \( A \) is a nonempty finite set of nonzero elements of \( V \), and if the number of elements of \( A \) is less than or equal to the number of elements in \( k \), then it follows that the union of \( V'_a \), \( a \in A \), is a proper subset of \( V' \). This means that there is a linear functional \( \lambda \) on \( V \) such that

\[
\lambda(a) \neq 0
\]

for every \( a \in A \).

### 18.13 Centralizers and diagonalizability

Let \( k \) be a field, and let \( (A, [\cdot, \cdot]_A) \) be a finite-dimensional Lie algebra over \( k \). Also let \( B \) be a nonzero Lie subalgebra of \( A \) such that every element of \( B \) is ad-diagonalizable as an element of \( A \). This implies that \( B \) is commutative as a Lie algebra over \( k \), as in Section 17.1. Let \( B' \) be the dual of \( B \), as a vector space over \( k \), as usual. If \( \alpha \in B' \), then we let \( A_\alpha \) be the set of \( x \in A \) such that \( \text{ad}_\alpha(x) = [w, x]_A = \alpha(w)x \) for every \( w \in B \), as before.

Thus \( A_\alpha \) is the same as the centralizer \( C_A(B) \) of \( B \) in \( A \). Of course, \( B \subseteq C_A(B) \), because \( B \) is commutative as a Lie algebra. Let \( \Phi_B \) be the set of
\( \alpha \in B' \) such that \( \alpha \neq 0 \) and \( A_{\alpha} \neq \{0\} \), as in Section 17.2. Remember that \( A \) corresponds to the direct sum of \( A_{\alpha}, \alpha \in \Phi_B \cup \{0\} \), as before.

Let \( w \in B \) be given, and let \( C_A(w) = C_A(\{w\}) \) be the centralizer of \( w \) in \( A \), which is the set of \( x \in A \) such that \([w, x]_A = 0\). This corresponds to the direct sum of \( A_{\alpha} \), where \( \alpha \in \Phi_B \cup \{0\} \) satisfies

\[
\alpha(w) = 0.
\]

Of course, this holds automatically when \( \alpha = 0 \). If

\[
\alpha(w) \neq 0
\]

for every \( \alpha \in \Phi_B \), then we get that

\[
C_A(w) = A_0 = C_A(B).
\]

This is related to the second part of Exercise 7 on p41 of [13].
Chapter 19

Root systems

19.1 Symmetries and reflections

Let $V$ be a finite-dimensional vector space over the real numbers, and let $\alpha$ be a nonzero element of $V$. A one-to-one linear mapping $\sigma$ from $V$ onto itself is said to be a symmetry with vector $\alpha$ if it satisfies the following two conditions. First,

$$\sigma(\alpha) = -\alpha.$$  

(19.1.1)

Second, the linear subspace

$$H = \{v \in V : \sigma(v) = v\}$$  

(19.1.2)

of $V$ is a hyperplane, which is to say that it has codimension one in $V$. Let

$$L_\alpha = \{t \alpha : t \in \mathbb{R}\}$$  

(19.1.3)

be the line in $V$ passing through $\alpha$. Note that

$$L_\alpha + H = V, \quad L_\alpha \cap H = \{0\},$$  

(19.1.4)

so that $L_\alpha$ and $H$ are complementary linear subspaces of $V$. It is easy to see that $\sigma$ is uniquely determined by $L_\alpha$ and $H$. One can also get a symmetry on $V$ from any line and complementary hyperplane in $V$, where the symmetry is the same as multiplication by $-1$ on the line, and the identity mapping on the hyperplane.

Let $V'$ be the dual space of linear functionals on $V$, as a vector space over $\mathbb{R}$, as usual. In this situation, there is a unique $\lambda_\alpha \in V'$ such that

$$\lambda_\alpha(\alpha) = 2$$  

(19.1.5)

and

$$\lambda_\alpha(v) = 0 \quad \text{for every } v \in H.$$  

(19.1.6)
Observe that
\[(19.1.7) \quad \sigma(v) = v - \lambda_\alpha(v) \alpha\]
for every \(v \in V\). Conversely, if \(\alpha \in V\) and \(\lambda_\alpha \in V'\) satisfy \((19.1.5)\), then
\[(19.1.7)\] defines a symmetry on \(V\) with vector \(\alpha\), for which the kernel of \(\lambda_\alpha\) is the fixed-point set. These remarks correspond to the discussion on p24 of [23].

Suppose now that \((v, w)\) is an inner product on \(V\). Let \(\sigma\) be a symmetry on \(V\) with vector \(\alpha\) and fixed hyperplane \(H\) as in \((19.1.2)\) again. Suppose that \(\alpha\) is orthogonal to \(H\), so that
\[(19.1.8) \quad (v, \alpha) = 0\]
for every \(v \in H\). Let us say that \(\sigma\) is a reflection on \(V\) with respect to \((\cdot, \cdot)\) in this case, as on p42 of [13]. One can check that \(\sigma\) is an orthogonal transformation on \(V\) with respect to \((\cdot, \cdot)\) under these conditions, which is to say that
\[(19.1.9) \quad (\sigma(v), \sigma(w)) = (v, w)\]
for every \(v, w \in V\).

Conversely, suppose that \(\sigma\) is a symmetry on \(V\) with vector \(\alpha\), and that \(\sigma\) is an orthogonal transformation on \(V\) with respect to \((\cdot, \cdot)\). If \(v\) is an element of the hyperplane \(H\) in \((19.1.2)\), then
\[(19.1.10) \quad (v, \alpha) = (\sigma(v), \sigma(\alpha)) = -(v, \alpha).\]
This implies \((19.1.8)\), so that \(\sigma\) is a reflection on \(V\). This is related to some of the remarks on p28 of [23].

Let \(\alpha \in V\), \(\alpha \neq 0\), be given, and let
\[(19.1.11) \quad H_\alpha = \{v \in V : (v, \alpha) = 0\}\]
be the hyperplane in \(V\) orthogonal to \(\alpha\). Put
\[(19.1.12) \quad \mu_\alpha(v) = 2(v, \alpha)(\alpha, \alpha)^{-1}\]
for every \(v \in V\), which defines a linear functional on \(V\). By construction,
\[(19.1.13) \quad \mu_\alpha(\alpha) = 2,\]
and \((19.1.11)\) is the kernel of \(\mu_\alpha\). The reflection associated to \(\alpha\) on \(V\) is given by
\[(19.1.14) \quad \sigma(v) = v - \mu_\alpha(v) \alpha = v - 2(v, \alpha)(\alpha, \alpha)^{-1} \alpha\]
for every \(v \in V\), as in \((19.1.7)\). This corresponds to some of the remarks on p42 of [13], and p28 of [23].

### 19.2 Invariant inner products

Let \(V\) be a vector space over the real numbers of positive finite dimension, and let \(G\) be a finite subgroup of the general linear group \(GL(V)\) of invertible
linear mappings on \( V \). Under these conditions, it is well known that there is an inner product \((v, w)\) on \( V \) that is invariant under \( G \), so that the elements of \( G \) are orthogonal transformations on \( V \) with respect to \((v, w)\). More precisely, if \((v, w)_0\) is any inner product on \( V \), then one can verify that

\[
(v, w) = \sum_{g \in G} (g(v), g(w))_0
\]

is an inner product on \( V \) that is invariant under the elements of \( G \).

Let \( A \) be a nonempty subset of \( V \), and let \( G_A \) be the collection of one-to-one linear mappings \( T \) from \( V \) onto itself such that

\[
T(A) = A.
\]

Of course, \( G_A \) is a subgroup of \( GL(V) \). The mapping from \( T \in G_A \) to the restriction of \( T \) to \( A \) defines a group homomorphism from \( G_A \) into the group of one-to-one mappings from \( A \) onto itself. Suppose that the linear span of \( A \) in \( V \) is equal to \( V \). If \( T \in G_A \) is equal to the identity mapping on \( A \), then it follows that \( T \) is the identity mapping on \( V \). If \( A \) has only finitely many elements, then \( G_A \) has only finitely many elements too. This implies that there is an inner product on \( V \) that is invariant under \( G_A \), as in the preceding paragraph.

Let \( \alpha \) be a nonzero element of \( V \), and let \( A \) be a finite subset of \( V \) whose linear span in \( V \) is equal to \( V \) again. Under these conditions, there is at most one symmetry \( \sigma \) on \( V \) with vector \( \alpha \) that maps \( A \) onto itself, as in the lemma on the top of p25 in [23]. To see this, let \((v, w)\) be an inner product on \( V \) that is invariant under \( G_A \), as in the preceding paragraph. Thus \( \sigma \) is an orthogonal transformation on \( V \) with respect to \((v, w)\), because \( \sigma \in G_A \), by hypothesis. It follows that \( \sigma \) is the reflection on \( V \) associated to \( \alpha \) with respect to \((v, w)\), as in the previous section.

Another version of this lemma is stated on p42 of [13], in terms of reflections. In this version, \( V \) is equipped with an inner product \((v, w)_V\), and \( A \) is a finite subset of \( V \) whose linear span in \( V \) is equal to \( V \) again. Let \( \sigma \) be a symmetry on \( V \) with vector \( \alpha \in A \), which is not necessarily a reflection on \( V \) with respect to \((v, w)_V\). Suppose that \( \sigma(A) = A \), and that the reflection \( \sigma_\alpha \) on \( V \) associated to \( \alpha \) with respect to \((v, w)_V\) maps \( A \) onto itself. Under these conditions, \( \sigma = \sigma_\alpha \).

This follows from the previous version, because \( \sigma_\alpha \) is a symmetry on \( V \) with vector \( \alpha \). Note however that \((v, w)_V\) is not necessarily invariant under \( G_A \). Of course, any symmetry on \( V \) with vector \( \alpha \) is the reflection associated to \( \alpha \) with respect to some inner product on \( V \).

### 19.3 Defining root systems

Let \( V \) be a vector space over the real numbers of positive finite dimension. A subset \( A \) of \( V \) is said to be a root system in \( V \) if it satisfies the following three conditions, as in Definition 1 on p25 of [23]. First, \( A \) has only finitely many
elements, the linear span of $A$ in $V$ is equal to $V$, and $0 \not\in A$. Second, for each $\alpha \in A$, there is a symmetry $\sigma_\alpha$ on $V$ with vector $\alpha$ such that

$$\sigma_\alpha(A) = A.$$  

(19.3.1)

Note that $\sigma_\alpha$ is uniquely determined by these properties, as in the previous section. Third, for every $\alpha, \beta \in A$,

$$\sigma_\alpha(\beta) - \beta \text{ is an integer multiple of } \alpha.$$  

(19.3.2)

The dimension of $V$ may be called the rank of $A$, and the elements of $A$ may be called roots of $V$ with respect to $A$.

If $\alpha \in A$, then there is a unique linear functional $\lambda_\alpha$ on $V$ such that $\lambda_\alpha(\alpha) = 2$ and $\lambda_\alpha = 0$ on the hyperplane in $V$ consisting of vectors fixed by $\sigma_\alpha$, as in Section 19.1. Remember that

$$\sigma_\alpha(v) = v - \lambda_\alpha(v) \alpha$$  

(19.3.3)

for every $v \in V$, as in (19.1.7). Using this, the third condition in the preceding paragraph can be reformulated as saying that for every $\alpha, \beta \in A$,

$$\lambda_\alpha(\beta) \in \mathbb{Z}.$$  

(19.3.4)

If $\alpha \in A$, then we also have that

$$\sigma_\alpha(\beta) - \beta \text{ is an integer multiple of } \alpha.$$  

(19.3.5)

by (19.3.1) and the fact that $s_\alpha(\alpha) = -\alpha$.

A root system $A$ in $V$ is said to be reduced if for every $\alpha \in A$, the only elements of $A$ that are proportional to $\alpha$ are $\alpha$ and $-\alpha$, as in Definition 2 on p25 of [23]. Otherwise, suppose that $\alpha \in A$ and $t \alpha \in A$ for some $t \in \mathbb{R}$ with $t \neq \pm 1$. Of course, $t \neq 0$, because $0 \not\in A$, and we may as well suppose that $|t| < 1$, since otherwise we could interchange the roles of $\alpha$ and $t \alpha$. Using (19.3.4), we get that $2t \in \mathbb{Z}$, because $\lambda_\alpha(\alpha) = 2$. This means that $t = \pm 1/2$.

In the definition of a root system on p42 of [13], one supposes that $V$ is already equipped with an inner product $(v, w)$, and that the symmetries $\sigma_\alpha$ are reflections with respect to this inner product. This means that

$$\lambda_\alpha(v) = 2(v, \alpha)(\alpha, \alpha)^{-1}$$  

(19.3.6)

for every $\alpha \in A$ and $v \in V$, as in Section 19.1. In this situation, (19.3.4) is the same as saying that

$$2(\beta, \alpha)(\alpha, \alpha)^{-1} \in \mathbb{Z}$$  

(19.3.7)

for every $\alpha, \beta \in A$. The condition that the root system be reduced is also included in the definition, as discussed further on p43 of [13]. If $A$ is a root system in $V$ as before, then one can find a compatible inner product on $V$, as in the next section.

Of course, we have seen properties like these in connection with Lie algebras, in Chapter 17. This corresponds to the remarks following the theorem on p40 of [13], and part (a) of Theorem 2 on p43 of [23]. This will be discussed further in Section 22.1. Some classical examples of root systems will be discussed starting in Section 20.10.
19.4 The Weyl group

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $A$ be a root system in $V$. The Weyl group of $A$ is defined to be the subgroup $W$ of $GL(V)$ generated by the symmetries $\sigma_\alpha$, $\alpha \in A$, as in Definition 3 on p27 of [23]. Let $G_A$ be the subgroup of $GL(V)$ consisting of invertible linear mappings $T$ on $V$ that map $A$ onto itself, as in Section 19.2. Thus

$$W \subseteq G_A,$$

(19.4.1)

because $\sigma_\alpha \in G_A$ for every $\alpha \in A$, by definition of a root system. Remember that $G_A$ has only finitely many elements in this situation, because $A$ is a finite subset of $V$ that spans $V$.

More precisely, $W$ is a normal subgroup of $G_A$, as mentioned on p27 of [23]. To see this, let $\sigma_\alpha$ and $T \in G_A$ be given. It is easy to see that $T \circ \sigma_\alpha \circ T^{-1}$ is a symmetry on $V$ with vector $T(\alpha)$. We also have that

$$T \circ \sigma_\alpha \circ T^{-1}(A) = A,$$

(19.4.2)

because $\sigma_\alpha$ and $T$ both map $A$ onto itself. This implies that

$$T \circ \sigma_\alpha \circ T^{-1} = \sigma_{T(\alpha)},$$

(19.4.3)

because $T(\alpha) \in A$, as in the previous section. It follows that $W$ is invariant under conjugation by $T$, as desired. Note that (19.4.3) corresponds to the first part of the lemma on p43 of [13], just after the definition of the Weyl group.

Let $\alpha \in A$ be given again, and let $\lambda_\alpha \in V'$ be as in the previous section, so that $\sigma_\alpha$ can be expressed as in (19.3.3). If $T \in GL(V)$, then it follows that

$$(T \circ \sigma_\alpha \circ T^{-1})(v) = v - \lambda_\alpha(T^{-1}(v))T(\alpha)$$

(19.4.4)

for every $v \in V$. Suppose that $T \in G_A$, and let $\lambda_{T(\alpha)}$ be the linear functional on $V$ corresponding to $T(\alpha) \in A$ in the same way. Thus

$$\sigma_{T(\alpha)}(v) = v - \lambda_{T(\alpha)}(v)T(\alpha)$$

(19.4.5)

for every $v \in V$, as before. This means that

$$\lambda_{T(\alpha)}(v) = \lambda_\alpha(T^{-1}(v))$$

(19.4.6)

for every $v \in V$, because of (19.4.3).

Because the Weyl group $W$ has only finitely many elements, one can find an inner product on $V$ that is invariant under $W$ as in Section 19.2. This corresponds to Proposition 1 on p27 of [23]. The choice of such an inner product is sometimes included in the definition of a root system, as in [13], and mentioned on p28 of [24].

Let $(v, w)$ be an inner product on $V$ that is invariant under $W$. If $\alpha \in A$, then $\lambda_\alpha \in V'$ can be expressed in terms of $(\cdot, \cdot)$ as in (19.3.6). Similarly, if $T \in G_A$, then $T(\alpha) \in A$, and

$$\lambda_{T(\alpha)}(v) = 2(v, T(\alpha))(T(\alpha), T(\alpha))^{-1}$$

(19.4.7)
for every \( v \in V \). Combining this with (19.4.6), we get that
\[
(v, T(\alpha))(T(\alpha), T(\alpha))^{-1} = (T^{-1}(v), \alpha) (\alpha, \alpha)^{-1}
\]
for every \( v \in V \). Equivalently, this means that
\[
(T(\alpha), T(\alpha))(T(\alpha), T(\alpha))^{-1} = (u, \alpha) (\alpha, \alpha)^{-1}
\]
for every \( u \in V \). This corresponds to the second part of the lemma on p43 of [13]. Note that this holds automatically when \((\cdot, \cdot)\) is invariant under \( G_A \).

19.5 Isomorphic root systems

Let \( V \) and \( \tilde{V} \) be vector spaces over the real numbers with the same positive finite dimension, and let \( A \) and \( \tilde{A} \) be root systems in \( V \) and \( \tilde{V} \), respectively. Let us say that a one-to-one linear mapping \( T \) from \( V \) onto \( \tilde{V} \) is an isomorphism between these root systems if
\[
T(A) = \tilde{A}.
\]
In particular, an invertible linear mapping on \( V \) is an automorphism of the root system \( A \) if it maps \( A \) onto itself, as on p27 of [23].

Let \( T : V \to \tilde{V} \) be an isomorphism between the root systems \( A \) and \( \tilde{A} \), and let \( \alpha \in A \) be given. If \( \sigma_\alpha \) is the corresponding symmetry on \( V \) with vector \( \alpha \), then it is easy to see that \( T \circ \sigma_\alpha \circ T^{-1} \) is a symmetry on \( \tilde{V} \) with vector \( T(\alpha) \). Observe that
\[
(T \circ \sigma_\alpha \circ T^{-1})(\tilde{A}) = \tilde{A},
\]
because of (19.5.1) and the fact that \( \sigma_\alpha(A) = A \). It follows that
\[
T \circ \sigma_\alpha \circ T^{-1} = \tilde{\sigma}_{T(\alpha)},
\]
where the right side is the symmetry on \( \tilde{V} \) corresponding to \( T(\alpha) \in \tilde{A} \) as in Section 19.3.

Of course, if \( \tau \) is a one-to-one linear mapping from \( V \) onto itself, then
\[
T \circ \tau \circ T^{-1}
\]
is a one-to-one linear mapping from \( \tilde{V} \) onto itself. More precisely,
\[
\tau \mapsto T \circ \tau \circ T^{-1}
\]
defines a group isomorphism from \( GL(V) \) onto \( GL(\tilde{V}) \). This isomorphism sends the Weyl group \( W \) of \( A \) onto the Weyl group \( \tilde{W} \) of \( \tilde{A} \), because of (19.5.3).

Let \( \lambda_\alpha \) be the linear functional on \( V \) corresponding to \( \sigma_\alpha \) as before, so that \( \sigma_\alpha \) can be expressed as in (19.3.3). Thus
\[
(T \circ \sigma_\alpha \circ T^{-1})(v) = \tilde{v} - \lambda_\alpha(T^{-1}(v)) T(\alpha)
\]
for every \( \tilde{v} \in \tilde{V} \). Let \( \tilde{\lambda}_{T(\alpha)} \) be the linear functional on \( \tilde{V} \) that corresponds to \( T(\alpha) \in \tilde{A} \) in the same way, so that

\[
\tilde{\sigma}_{T(\alpha)}(\tilde{v}) = \tilde{v} - \tilde{\lambda}_{T(\alpha)}(\tilde{v}) T(\alpha)
\]

for every \( \tilde{v} \in \tilde{V} \). It follows that

\[
\tilde{\lambda}_{T(\alpha)}(\tilde{v}) = \lambda_\alpha(T^{-1}(\tilde{v}))
\]

for every \( \tilde{v} \in \tilde{V} \), by (19.5.3). Equivalently,

\[
\tilde{\lambda}_{T(\alpha)}(T(v)) = \lambda_\alpha(v)
\]

for every \( v \in V \).

Suppose that \((\cdot, \cdot)_V\) and \((\cdot, \cdot)_{\tilde{V}}\) are inner products on \( V \) and \( \tilde{V} \) that are compatible with the root systems \( A \) and \( \tilde{A} \), respectively, in the sense that \( \sigma_\alpha \) and \( \tilde{\sigma}_\alpha \) are reflections on \( V \) and \( \tilde{V} \) with respect to \((\cdot, \cdot)_V\) and \((\cdot, \cdot)_{\tilde{V}}\) for every \( \alpha \in A \) and \( \tilde{\alpha} \in \tilde{A} \), respectively. Thus

\[
\lambda_\alpha(v) = 2(v, \alpha)_V (\alpha, \alpha)_V^{-1}
\]

for every \( \alpha \in A \) and \( v \in V \), as in (19.3.6), and similarly

\[
\tilde{\lambda}_\alpha(\tilde{v}) = 2(\tilde{v}, \tilde{\alpha})_{\tilde{V}} (\tilde{\alpha}, \tilde{\alpha})_{\tilde{V}}^{-1}
\]

for every \( \tilde{\alpha} \in \tilde{A} \) and \( \tilde{v} \in \tilde{V} \). In this case, (19.5.9) is the same as saying that

\[
(T(v), T(\alpha))_{\tilde{V}} (T(\alpha), T(\alpha))_{\tilde{V}}^{-1} = (v, \alpha)_V (\alpha, \alpha)_V^{-1}
\]

for every \( \alpha \in A \) and \( v \in V \). This is related to some of the remarks on p43 of [13]. In particular, (19.5.12) holds automatically when

\[
(T(u), T(v))_{\tilde{V}} = (u, v)_V
\]

for every \( u, v \in V \).

19.6 Duality and symmetries

Let \( V \) be a vector space over the real numbers of positive finite dimension, and let \( V' \) be the dual space of all linear functionals on \( V \). Remember that \( V' \) has the same dimension as \( V \). If \( v \in V \) and \( \mu \in V' \), then put

\[
L_v(\mu) = \mu(v).
\]

This defines \( L_v \) as a linear functional on \( V' \), so that \( L_v \) is an element of the dual space \( V'' = (V')' \) of \( V' \). More precisely, \( v \mapsto L_v \) is a one-to-one linear mapping from \( V \) onto \( V'' \).
If $T$ is a linear mapping from $V$ into itself and $\mu \in V'$, then
\begin{equation}
T'(\mu) = \mu \circ T
\end{equation}
defines a linear functional on $V$ too. Equivalently,
\begin{equation}
(T'(\mu))(v) = \mu(T(v))
\end{equation}
for every $v \in V$. This defines $T'$ as a linear mapping from $V'$ into itself, which is the dual linear mapping associated to $T$. If $T_1, T_2$ are linear mappings from $V$ into itself, then
\begin{equation}
(T_1 \circ T_2)' = T_2' \circ T_1'
\end{equation}
as linear mappings from $V'$ into itself. Note that
\begin{equation}
(I_V')' = I_{V'}
\end{equation}
where $I_V$ and $I_{V'}$ are the identity mappings on $V$ and $V'$, respectively.

If $T$ is a linear mapping from $V$ into itself again, then $T''$ is a linear mapping from $V'$ into itself, which leads to a dual mapping $T''' = (T')'$ from $V''$ into itself. Thus, if $L \in V''$, then
\begin{equation}
T'''(L) = L \circ T',
\end{equation}
as a linear functional on $V'$. This means that
\begin{equation}
(T'''(L))(\mu) = L(T'(\mu)) = L(\mu \circ T)
\end{equation}
for every $\mu \in V'$. If $v \in V$ and $L_v \in V''$ is as in (19.6.1), then it follows that
\begin{equation}
(T'''(L_v))(\mu) = L_v(\mu \circ T) = \mu(T(v)) = L_{T(v)}(\mu)
\end{equation}
for every $\lambda \in V'$. This means that
\begin{equation}
T'''(L_v) = L_{T(v)}
\end{equation}
for every $v \in V$.

Let $z \in V$ and $\tau \in V'$ be given, so that
\begin{equation}
T_{z,\tau}(v) = \tau(v) z
\end{equation}
deﬁnes a linear mapping from $V$ into itself. If $\mu \in V'$, then
\begin{equation}
(T'_{z,\tau}(\mu))(v) = \mu(T_{z,\tau}(v)) = \tau(v) \mu(z) = \tau(v) L_z(\mu)
\end{equation}
for every $v \in V$. Equivalently, this means that
\begin{equation}
T'_{z,\tau}(\mu) = L_z(\mu) \tau
\end{equation}
for every $\mu \in V'$. 
19.7. ADJOINTS AND REFLECTIONS

Let $\sigma$ be a symmetry on $V$ with vector $\alpha \in V$. Thus there is a linear functional $\lambda_\alpha$ on $V$ such that $\lambda_\alpha(\alpha) = 2$ and $\sigma(v) = v - \lambda_\alpha(v) \alpha$ for every $v \in V$, as in Section 19.1. This is the same as saying that

$$\sigma = I_V - T_{\alpha, \lambda_\alpha},$$

where $T_{\alpha, \lambda_\alpha}$ is as in (19.6.10). The corresponding dual linear mapping is given by

$$\sigma'(\mu) = \mu - L_\alpha(\mu) \lambda_\alpha$$

for every $\mu \in V'$, by (19.6.5) and (19.6.12). Note that

$$L_\alpha(\lambda_\alpha) = \lambda_\alpha(\alpha) = 2.$$

It follows that $\sigma'$ is a symmetry on $V'$ with vector $\lambda_\alpha$, as in Section 19.1. Alternatively, one can check more directly that

$$\sigma'(\lambda_\alpha) = \lambda_\alpha \circ \sigma = -\lambda_\alpha,$$

because $\sigma(\alpha) = -\alpha$ and the kernel of $\lambda_\alpha$ is the fixed-point set of $\sigma$. If $\mu \in V'$ satisfies $\mu(\alpha) = 0$, then one can verify that

$$\sigma'(\mu) = \mu \circ \sigma = \mu,$$

because $\sigma$ is equal to the identity on a hyperplane in $V$ complementary to the line passing through $\alpha$.

19.7 Adjoints and reflections

Let $V$ be a vector space over the real numbers of positive finite dimension again, and suppose that $(v, w)$ is an inner product on $V$. If $T$ is a linear mapping from $V$ into itself, then there is a unique adjoint linear mapping $T^*$ from $V$ into itself such that

$$(T(v), w) = (v, T^*(w))$$

for every $v, w \in V$, as before. Remember that

$$(T_1 \circ T_2)^* = T_2^* \circ T_1^*$$

for all linear mappings $T_1$ and $T_2$ from $V$ into itself, and that $I_V^* = I_V$. If $T$ is any linear mapping from $V$ into itself, then $(T^*)^* = T$. We also have that $T$ is an orthogonal transformation on $V$ with respect to $(\cdot, \cdot)$, in the sense that

$$(T(v), T(w)) = (v, w)$$

for every $v, w \in V$, if and only if $T$ is invertible on $V$, with $T^{-1} = T^*$. If $u \in V$, then

$$\mu_u(v) = (v, u)$$
Chapter 19. Root Systems

defines a linear functional on $V$, and $u \mapsto \mu_u$ is a one-to-one linear mapping from $V$ onto $V'$. If $T$ is a linear mapping from $V$ into itself, then it is easy to see that

$$T'(\mu_u) = \mu_{T^*u}$$

for every $u \in V$. Let $u, z \in V$ be given, and put

$$T_{z,u}(v) = (v, u) z$$

for every $v \in V$. This defines a linear mapping from $V$ into itself, which corresponds to (19.6.10), with $\tau = \mu_u$. Observe that

$$(T_{z,u}(v), w) = (v, u) (z, w) = (v, Tu_z(w))$$

for every $v, w \in V$, so that

$$(T_{z,u})^* = T_{u,z}.$$ 

Let $\alpha \in V$ with $\alpha \neq 0$ be given, and put

$$\hat{\alpha} = 2 \alpha (\alpha, \alpha)^{-1}.$$ 

Thus

$$\mu_{\alpha^*}(v) = (v, \hat{\alpha}) = 2 (v, \alpha) (\alpha, \alpha)^{-1}$$

for every $v \in V$, in the notation of (19.7.4). Note that

$$\mu_\alpha(\alpha) = 2,$$

and that

$$\sigma_\alpha(v) = v - \mu_\alpha^-(v) \alpha = v - 2 (v, \alpha) (\alpha, \alpha)^{-1} \alpha$$

is the reflection on $V$ associated to $\alpha$ with respect to $(\cdot, \cdot)$, as in Section 19.1. It is easy to see that

$$\sigma_\alpha^* = \sigma_\alpha,$$

using (19.7.8). This also corresponds to the fact that $\sigma_\alpha$ is an orthogonal transformation on $V$, and its own inverse.

Observe that $\sigma_\alpha$ may be considered as the reflection on $V$ associated to $\hat{\alpha}$ with respect to $(\cdot, \cdot)$ as well, so that

$$\sigma_\alpha = \sigma_{\hat{\alpha}}.$$ 

Indeed, $\sigma_\alpha$ only depends on the line in $V$ passing through $\alpha$. Put

$$(\hat{\alpha}) = 2 \alpha (\alpha, \alpha)^{-1},$$

as in (19.7.9). Clearly

$$(\hat{\alpha}, \hat{\alpha}) = 4 (\alpha, \alpha) (\alpha, \alpha)^{-2} = 4 (\alpha, \alpha)^{-1},$$

so that

$$(\hat{\alpha}) = 4 \alpha (\alpha, \alpha)^{-1} (4 (\alpha, \alpha)^{-1})^{-1} = \alpha.$$
Using (19.7.12), we have that
\[ \sigma_\alpha(v) = v - \mu_\alpha(v) \hat{\alpha} = v - \mu_{\beta}(v) \hat{\alpha} \]
for every \( v \in V \), which is another way to look at (19.7.14).

Let \( \beta \in V \) with \( \beta \neq 0 \) be given, and put \( \hat{\beta} = 2 \beta \langle \beta, \beta \rangle^{-1} \), as in (19.7.9). If 
\( T \) is an orthogonal transformation on \( V \) with respect to \( \langle \cdot, \cdot \rangle \), then
\[ T(\hat{\beta}) = 2T(\beta) \langle \beta, \beta \rangle^{-1} = 2T(\beta) (T(\beta), T(\beta))^{-1} = T(\hat{\beta}), \]
where \( T(\hat{\beta}) \) is defined as in (19.7.9) too.

## 19.8 Inverse roots

Let \( V \) be a vector space over the real numbers of positive finite dimension, and let \( A \) be a root system in \( V \). If \( \alpha \in A \), then there is a unique symmetry \( \sigma_\alpha \) on \( V \) with vector \( \alpha \) that maps \( A \) onto itself, as in Section 19.3. Let \( \lambda_\alpha \) be the linear functional on \( V \) such that \( \lambda_\alpha(\alpha) = 2 \) and \( \lambda_\alpha = 0 \) on the hyperplane of vectors in \( V \) fixed by \( \alpha \), so that
\[ \sigma_\alpha(v) = v - \lambda_\alpha(v) \alpha \]
for every \( v \in V \), as before. As on p25 of [23], \( \lambda_\alpha \) may be called the inverse root associated to \( \alpha \) in the dual space \( V' \) of \( V \). Let
\[ A' = \{ \lambda_\alpha : \alpha \in A \} \]
be the set of these inverse roots, which is a finite subset of \( V' \) that does not contain \( 0 \).

Let \( (v, w) \) be an inner product on \( V \) that is invariant under the Weyl group \( W \) of \( A \), as in Section 19.4. Thus \( \sigma_\alpha \) is the reflection on \( V \) associated to \( \alpha \) with respect to \( (v, w) \), as in Section 19.1. In this situation,
\[ \lambda_\alpha(v) = \mu_\alpha(v) = (v, \hat{\alpha}) = 2 (v, \alpha) (\alpha, \alpha)^{-1} \]
for every \( v \in V \), using the notation in (19.7.4) and (19.7.9). Put
\[ \hat{A} = \{ \hat{\alpha} : \alpha \in A \} = \{ 2 \alpha (\alpha, \alpha)^{-1} : \alpha \in A \}, \]
which is a finite subset of \( V \) that does not contain \( 0 \). It is easy to see that the linear span of (19.8.4) in \( V \) is equal to \( V \), because of the corresponding property of \( A \). Note that (19.8.4) corresponds to (19.8.2) under the isomorphism \( u \mapsto \mu_u \) from \( V \) onto \( V' \) associated to the inner product. It follows that the linear span of \( A' \) in \( V' \) is equal to \( V' \).

If \( \beta \in V \), \( \beta \neq 0 \), and \( \hat{\beta} \) is as in (19.7.9), then
\[ \sigma_\alpha(\hat{\beta}) = (\sigma_\alpha(\beta)), \]
where \((\sigma_\alpha(\beta))\) is defined as in (19.7.9) too. This follows from (19.7.19), because \(\sigma_\alpha\) is an orthogonal transformation on \(V\) with respect to \((\cdot, \cdot)\). This implies that

\[
\sigma_\alpha(\hat{A}) = \hat{A},
\]

because \(\sigma_\alpha(A) = A\). This is the same as saying that

\[
\sigma_\alpha^*(\hat{A}) = \hat{A},
\]

because \(\sigma_\alpha^* = \sigma_\alpha\), as in (19.7.14). Using (19.8.6), we get that

\[
\sigma'_\alpha(A') = A',
\]

where \(\sigma'_\alpha\) is the dual linear mapping on \(V'\) associated to \(\sigma_\alpha\). This also uses the correspondence between (19.8.2) and (19.8.4) mentioned earlier, the correspondence between adjoints and dual linear mappings in (19.7.5), and the self-adjointness of \(\sigma_\alpha\), as in (19.7.13). Remember that \(\sigma'_\alpha\) is a symmetry on \(V'\) with vector \(\lambda_\alpha\), as in Section 19.6.

If \(\mu \in V'\), then

\[
\sigma'_\alpha(\mu) = \mu - L_\alpha(\mu) \lambda_\alpha = \mu - \mu(\alpha) \lambda_\alpha,
\]

as in (19.6.1) and (19.6.14). Let \(\beta \in A\) be given, and let \(\lambda_\beta\) be the corresponding element of (19.8.2). We would like to check that

\[
\sigma'_\alpha(\lambda_\beta) - \lambda_\beta \text{ is an integer multiple of } \lambda_\alpha.
\]

This is the same as saying that

\[
\lambda_\beta(\alpha) \in \mathbb{Z},
\]

by (19.8.9). This condition holds because \(A\) is a root system in \(V\), as in Section 19.3. This shows that \(A'\) is a root system in \(V'\), as in Proposition 2 on p28 of [23]. This is called the inverse system or dual system of \(A\) in \(V'\).

Alternatively, let us check that

\[
\sigma^*_\alpha(\hat{\beta}) - \hat{\beta} \text{ is an integer multiple of } \hat{\alpha}.
\]

This means that

\[
(\alpha, \hat{\beta}) \in \mathbb{Z},
\]

by (19.7.14) and (19.7.18). This condition holds because \(A\) is a root system in \(V\), as before. This shows that \(\hat{A}\) is a root system in \(V\), which may be called the inverse system or dual system of \(A\) in \(V\) with respect to \((\cdot, \cdot)\). This corresponds to the formulation on p43 of [13].

If \(T\) is an invertible linear mapping on \(V\), then the dual linear mapping \(T'\) is invertible on \(V'\), with

\[
(T')^{-1} = (T^{-1})'.
\]
It is well known and easy to see that the mapping from $T$ to (19.8.14) defines a group isomorphism from $GL(V)$ onto $CL(V')$. This isomorphism takes the Weyl group of $A$ onto the Weyl group of $A'$.

Similarly, if $T$ is an invertible linear mapping on $V$, then the adjoint $T^*$ of $T$ with respect to $(\cdot, \cdot)$ is invertible on $V$ as well, with

\[(T^*)^{-1} = (T^{-1})^*.
\]

The mapping from $T$ to (19.8.15) defines a group automorphism on $GL(V)$. The Weyl group of $A$ is the same as the Weyl group of $A$, because it is generated by the same reflections on $V$.

### 19.9 Angles between roots

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $A$ be a root system in $V$. If $\alpha \in A$, then let $\sigma_\alpha$ be the unique symmetry on $V$ with vector $\alpha$ that maps $A$ onto itself, as in Section 19.3. Also let $\lambda_\alpha$ be the linear functional on $V$ such that $\lambda_\alpha(\alpha) = 2$ and $\lambda_\alpha = 0$ on the hyperplane of vectors in $V$ fixed by $\sigma_\alpha$, as before. If $\beta \in A$ too, then put

\[(19.9.1) \quad n(\beta, \alpha) = \lambda_\alpha(\beta),
\]

as on p29 of [23]. Remember that $n(\beta, \alpha) \in \mathbb{Z}$, as in Section 19.3.

Let $(v, w)$ be an inner product on $V$ that is invariant under the Weyl group of $A$, as in Section 19.4. If $v \in V$, then put

\[(19.9.2) \quad \|v\| = (v, v)^{1/2},
\]

which is the norm on $V$ associated to $(\cdot, \cdot)$. Note that

\[(19.9.3) \quad n(\beta, \alpha) = 2 (\beta, \alpha) (\alpha, \alpha)^{-1}
\]

for every $\alpha, \beta \in A$, as in Section 19.3. Of course,

\[(19.9.4) \quad (\beta, \alpha) = \|\alpha\| \|\beta\| \cos \theta,
\]

where $\theta \in [0, \pi]$ is the angle between $\alpha$ and $\beta$ with respect to $(\cdot, \cdot)$. Thus

\[(19.9.5) \quad n(\beta, \alpha) = 2 \|\alpha\|^{-1} \|\beta\| \cos \theta.
\]

It follows that

\[(19.9.6) \quad n(\beta, \alpha) n(\alpha, \beta) = 4 \cos^2 \theta
\]

for every $\alpha, \beta \in A$. The left side is an integer, and the right side is a nonnegative real number less than or equal to 4. This means that $4 \cos^2 \theta$ can only take the values 0, 1, 2, 3, 4. Of course,

\[(19.9.7) \quad 4 \cos^2 \theta = 4
\]
CHAPTER 19. ROOT SYSTEMS

exactly when $\theta = 0$ or $\pi$, which means that $\alpha$ and $\beta$ are proportional in $V$. Let us suppose now that $\alpha$ and $\beta$ are not proportional in $V$, so that $0 < \theta < \pi$.

Similarly,

(19.9.8) \[ 4 \cos^2 \theta = 0 \]

exactly when $\theta = \pi/2$, in which case

(19.9.9) \[ n(\beta, \alpha) = n(\alpha, \beta) = 0. \]

Otherwise, if $\theta \neq \pi/2$, then $\cos \theta \neq 0$, $(\alpha, \beta) \neq 0$, and

(19.9.10) \[ n(\alpha, \beta)/n(\beta, \alpha) = ||\alpha||^2/||\beta||^2. \]

Remember that $\cos^2 \theta + \sin^2 \theta = 1$. Note that

(19.9.11) \[ 4 \cos^2 \theta = 2 \]

if and only if $\theta = \pi/4$ or $3\pi/4$. If $\theta = \pi/4$, then we either have that

(19.9.12) \[ n(\alpha, \beta) = 1, \quad n(\beta, \alpha) = 2, \quad \text{and} \quad ||\beta|| = \sqrt{2}||\alpha||, \]

or the analogous conditions with $\alpha$ and $\beta$ interchanged. If $\theta = 3\pi/4$, then we either have that

(19.9.13) \[ n(\alpha, \beta) = -1, \quad n(\beta, \alpha) = -2, \quad \text{and} \quad ||\beta|| = \sqrt{2}||\alpha||, \]

or the analogous conditions with $\alpha$ and $\beta$ interchanged.

We also have that

(19.9.14) \[ 4 \cos^2 \theta = 1 \]

if and only if $\theta = \pi/3$ or $2\pi/3$. More precisely,

(19.9.15) \[ n(\alpha, \beta) = n(\beta, \alpha) = 1 \]

when $\theta = \pi/3$, and

(19.9.16) \[ n(\alpha, \beta) = n(\beta, \alpha) = -1 \]

when $\theta = 2\pi/3$. In both cases,

(19.9.17) \[ ||\alpha|| = ||\beta||. \]

Similarly,

(19.9.18) \[ 4 \cos^2 \theta = 3 \]

if and only if $\theta = \pi/6$ or $5\pi/6$. If $\theta = \pi/6$, then we either have that

(19.9.19) \[ n(\alpha, \beta) = 1, \quad n(\beta, \alpha) = 3, \quad \text{and} \quad ||\beta|| = \sqrt{3}||\alpha||, \]

or the analogous conditions with $\alpha$ and $\beta$ interchanged. If $\theta = 5\pi/6$, then we either have that

(19.9.20) \[ n(\alpha, \beta) = -1, \quad n(\beta, \alpha) = -3, \quad \text{and} \quad ||\beta|| = \sqrt{3}||\alpha||, \]
or the analogous conditions with \( \alpha \) and \( \beta \) interchanged. This corresponds to remarks on p44 of [13], and p29 of [23].

If \( \alpha \) and \( \beta \) are not proportional in \( V \) and

\[
(19.9.21) \quad n(\alpha, \beta) > 0,
\]

then

\[
(19.9.22) \quad \alpha - \beta \in A,
\]

as in Proposition 3 on p29 of [23]. Note that (19.9.21) is the same as saying that \( \theta < \pi/2 \), which is symmetric in \( \alpha \) and \( \beta \). In this situation, we have that \( n(\alpha, \beta) = 1 \) or \( n(\beta, \alpha) = 1 \), as before. Suppose that \( n(\beta, \alpha) = 1 \), so that \( \lambda_\alpha(\beta) = 1 \), as in (19.9.1). In this case,

\[
(19.9.23) \quad \alpha - \beta = -(\beta - \lambda_\alpha(\beta) \alpha) = -\sigma_\alpha(\beta),
\]

by the usual expression for \( \sigma_\alpha \) in terms of \( \lambda_\alpha \). This implies (19.9.22), because \( \sigma_\alpha(\beta) \in A \), and \( A \) is symmetric about 0 in \( V \). If \( n(\alpha, \beta) = 1 \), then \( \alpha - \beta = \sigma_\beta(\alpha) \in A \).

Of course, (19.9.21) is equivalent to

\[
(19.9.24) \quad (\alpha, \beta) > 0,
\]

by (19.9.3). The lemma on p45 of [13] is stated in terms of (19.9.24), where it is also mentioned that

\[
(19.9.25) \quad \alpha + \beta \in A
\]

when \( \alpha \) and \( \beta \) are not proportional and

\[
(19.9.26) \quad (\alpha, \beta) < 0.
\]

This follows from the previous statement, applied to \( \beta \) in place of \(-\beta\).

### 19.10 A criterion for linear independence

Let \( V \) be a vector space over the real numbers, and let \((v, w)\) be an inner product on \( V \). Also let \( \tau \) be a linear functional on \( V \), and let \( B \) be a nonempty subset of \( V \). Suppose that

\[
(19.10.1) \quad \tau(\beta) > 0
\]

for every \( \beta \in B \), and that

\[
(19.10.2) \quad (\beta, \gamma) \leq 0
\]

for every \( \beta, \gamma \in B \) with \( \beta \neq \gamma \). Under these conditions, the elements of \( B \) are linearly independent in \( V \). This corresponds to Lemma 4 on p31 of [23], and Step (3) on p48 of [13].

To see this, observe that a linear relation between elements of \( B \) can be expressed as

\[
(19.10.3) \quad \sum_{\beta \in B_1} b_\beta \beta = \sum_{\gamma \in B_2} c_\gamma \gamma,
\]
where $B_1$ and $B_2$ are nonempty disjoint finite subsets of $B$, and the coefficients $b_\beta$ and $c_\gamma$ are nonnegative real numbers. Let $v \in V$ be the vector determined by (19.10.3), so that

$$ (v, v) = \sum_{\beta \in B_1} \sum_{\gamma \in B_2} b_\beta c_\gamma (\beta, \gamma). $$

(19.10.4)

It follows that $(v, v) \leq 0$, by (19.10.2). Of course, this means that $v = 0$.

Using this, we get that

$$ 0 = \tau(v) = \sum_{\beta \in B_1} b_\beta \tau(\beta). $$

(19.10.5)

This implies that $b_\beta = 0$ for every $\beta \in B_1$, because of (19.10.1). Similarly, $c_\gamma = 0$ for every $\gamma \in B_2$, as desired.

### 19.11 Bases for root systems

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $A$ be a root system in $V$. A subset $B$ of $A$ is said to be a base for $A$ if it satisfies the following two conditions. The first condition is that $B$ is a basis for $V$ as a vector space over $\mathbb{R}$. The second condition is that every $\alpha \in A$ can be expressed as

$$ \alpha = \sum_{\beta \in B} m_\beta \beta, $$

(19.11.1)

where $m_\beta \in \mathbb{Z}$ for every $\beta \in B$, and either $m_\beta \geq 0$ for every $\beta \in B$, or $m_\beta \leq 0$ for every $\beta \in B$. This is Definition 4 on p30 of [23], which corresponds to the definition of a base on p47 of [13]. A base for a root system may also be called a simple root system or a fundamental root system, as mentioned on p30 of [23]. Similarly, the elements of the base may be called simple roots.

Of course, the second condition in the preceding paragraph implies in particular that $A$ is contained in the linear span of $B$ in $V$. Remember that the linear span of $A$ in $V$ is equal to $V$, by definition of a root system. Thus the second condition implies that the linear span of $B$ in $V$ is equal to $V$ too.

Theorem 1 on p30 of [23] states that bases exist, which corresponds to the theorem at the top of p48 of [13]. More precisely, if $\tau$ is a linear functional on $V$, then put

$$ A_\tau^+ = \{\alpha \in A : \tau(\alpha) > 0\}. $$

(19.11.2)

Let us choose $\tau \in V'$ so that for every $\alpha \in A$,

$$ \tau(\alpha) \neq 0. $$

(19.11.3)

One can find such a $\tau \in V'$ as in Section 18.12. This implies that

$$ A = A_\tau^+ \cup (-A_\tau^+). $$

(19.11.4)
Let us say that \( \alpha \in A_+^\tau \) is decomposable if it can be expressed as the sum of two elements of \( A_+^\tau \), and indecomposable otherwise. The first part of Proposition 4 on p30 of [23] states that

\[ B_\tau = \{ \alpha \in A_+^\tau : \alpha \text{ is indecomposable} \} \]  

is a base for \( A \). This corresponds to the first part of the second theorem on p48 of [13].

The first step in the proof is Lemma 2 on p30 of [23], which corresponds to Step (1) on p48 of [13]. This states that every element of \( A_+^\tau \) can be expressed as a linear combination of elements of \( B_\tau \), with coefficients that are nonnegative integers. Otherwise, there would be an \( \alpha \in A_+^\tau \) that does not have this property, and for which \( \tau(\alpha) \) is as small as possible. In particular, \( \alpha \notin B_\tau \), so that \( \alpha \) is decomposable. Thus there are \( \alpha_1, \alpha_2 \in A_+^\tau \) whose sum is equal to \( \alpha \). This implies that

\[ \tau(\alpha) = \tau(\alpha_1) + \tau(\alpha_2), \]

and hence

\[ \tau(\alpha_1), \tau(\alpha_2) < \tau(\alpha), \]

because \( \tau(\alpha_1), \tau(\alpha_2) > 0 \). It follows that \( \alpha_1 \) and \( \alpha_2 \) can be expressed as linear combinations of elements of \( B_\tau \) with coefficients that are nonnegative integers, by the minimality of \( \tau(\alpha) \). This means that \( \alpha \) has the same property, which is a contradiction.

Let \( (v, w) \) be an inner product on \( V \) that is invariant under the Weyl group of \( A \), as in Section 19.4. If \( \beta_1 \) and \( \beta_2 \) are distinct elements of \( B_\tau \), then

\[ (\beta_1, \beta_2) \leq 0. \]  

This corresponds to Lemma 3 on p30 of [23], and Step (2) on p48 of [13]. Suppose for the sake of a contradiction that

\[ (\beta_1, \beta_2) > 0. \]  

If \( \beta_1 \) and \( \beta_2 \) are proportional in \( V \), then the proportionality constant is positive, by (19.11.9). This implies that one of \( \beta_1 \) and \( \beta_2 \) is equal to 2 times the other, as in Section 19.3, because \( \beta_1 \neq \beta_2 \), by hypothesis. However, this would contradict the fact that the elements of \( B_\tau \) are indecomposable. Otherwise, if \( \beta_1 \) and \( \beta_2 \) are not proportional in \( V \), then \( \gamma = \beta_1 - \beta_2 \in A \), as in Section 19.9. If \( \gamma \in A_+^\tau \), then \( \beta_1 = \beta_2 + \gamma \) is decomposable, contradicting the hypothesis that \( \beta_1 \in B_\tau \). Similarly, if \( -\gamma \in A_+^\tau \), then \( \beta_2 = \beta_1 + (-\gamma) \) is decomposable, contradicting the hypothesis that \( \beta_2 \in B_\tau \).

Using (19.11.8) and the fact that \( B_\tau \subseteq A_+^\tau \), by construction, we get that the elements of \( B_\tau \) are linearly independent in \( V \), as in the previous section. Remember that the elements of \( A_+^\tau \) can be expressed as linear combinations of elements of \( B_\tau \), with coefficients that are nonnegative integers. Thus the elements of \( -A_+^\tau \) can be expressed as linear combinations of elements of \( B_\tau \) with coefficients that are integers less than or equal to 0. This shows that \( B_\tau \)
satisfies the second condition in the definition of a base for $A$. In particular, the linear span of $B_\tau$ in $V$ is equal to $V$, as before. It follows that $B_\tau$ is a basis for $V$ as a vector space over $\mathbb{R}$. This implies that $B_\tau$ is a base for $A$, as on p31 of [23], and Step (4) on p48 of [13].

Conversely, suppose that $B$ is any base for $A$. Let $\tau$ be a linear functional on $V$ such that
\begin{equation}
\tau(\beta) > 0
\end{equation}
for every $\beta \in B$. It is easy to find such a $\tau \in V'$, because $B$ is a basis for $V$. Let $A^+$ be the set of $\alpha \in A$ that can be expressed as a linear combination of elements of $B$ with coefficients that are nonnegative integers. Observe that
\begin{equation}
A^+ \subseteq A^+_\tau,
\end{equation}
where $A^+_\tau$ is as in (19.11.2). This implies that
\begin{equation}
-A^+ \subseteq -A^+_\tau.
\end{equation}
It follows that
\begin{equation}
A^+ = A^+_\tau,
\end{equation}
because $A = A^+ \cup (-A^+)$, by hypothesis. We also get that (19.11.3) holds, so that the remarks in the previous paragraphs hold for $\tau$. Using (19.11.13), we obtain that the elements of $B$ are indecomposable in $A^+_\tau$, so that $B \subseteq B_\tau$. This implies that
\begin{equation}
B = B_\tau,
\end{equation}
because $B$ and $B_\tau$ have the same number of elements, which is the dimension of $V$. This corresponds to the second part of Proposition 4 on p30 of [23], and the second part of the second theorem on p48 of [13].

If $\beta_1$ and $\beta_2$ are distinct elements of $B$, then it is easy to see that their difference is not a root, using the second condition in the definition of a base for $A$. We also have that $\beta_1$ and $\beta_2$ are not proportional in $V$, because $B$ is a basis for $V$. This implies that (19.11.8) holds, as in Section 19.9. This corresponds to the lemma on p47 of [13]. This could also be obtained from (19.11.14) and the previous proof, but this argument is more direct.

19.12 Positive roots

Let $V$ be a vector space over the real numbers of positive finite dimension again, and let $A$ be a root system in $V$. Also let $B$ be a base for $A$, and let $A^+$ be the set of $\alpha \in A$ such that $\alpha$ can be expressed as a linear combination of elements of $B$ with coefficients that are nonnegative integers, as in the previous section. The elements of $A^+$ are called positive roots with respect to $B$, as on p31 of [23], and p47 of [13].

Let $\tau$ be the linear functional on $V$ such that
\begin{equation}
\tau(\beta) = 1
\end{equation}

19.12. POSITIVE ROOTS

for every \( \beta \in B \). If \( \gamma \in A^+ \), then

\[
\tau(\gamma) \in \mathbb{Z}_+.
\]

Let \((v, w)\) be an inner product on \( V \) that is invariant under the Weyl group of \( A \), as in Section 19.4. If \( \gamma \in A^+ \), then we would like to check that there is a \( \beta \in B \) such that

\[
\beta, \gamma > 0.
\]

Suppose for the sake of a contradiction that \((\beta, \gamma) \leq 0\) for every \( \beta \in B \). Remember that \((\beta_1, \beta_2) \leq 0\) for all \( \beta_1, \beta_2 \in B \) with \( \beta_1 \neq \beta_2 \), as in the previous section. Under these conditions, we get that \( B \cup \{\gamma\} \) is linearly independent in \( V \), as in Section 19.10, which is a contradiction.

Let \( \gamma \in A^+ \) be given, and put \( r = \tau(\gamma) \in \mathbb{Z}_+ \). Proposition 5 on p32 of [23] states that \( \gamma \) can be expressed as

\[
\gamma = \sum_{j=1}^{r} \alpha_j,
\]

where \( \alpha_j \in B \) for each \( j = 1, \ldots, r \), and

\[
\sum_{j=1}^{l} \alpha_j \in A
\]

for every \( l = 1, \ldots, r \). This corresponds to the corollary to Lemma A on p50 of [13]. If \( r = 1 \), then \( \gamma \in B \), and this is trivial. Otherwise, we can use induction on \( r \), as follows.

As before, there is a \( \beta \in B \) such that (19.12.3) holds. If \( \beta \) and \( \gamma \) are proportional in \( V \), then one can check that either \( \gamma = \beta \) or \( \gamma = 2\beta \), using remarks in Section 19.3. The statement in the preceding paragraph obviously holds in both cases, and so we suppose now that \( \beta \) and \( \gamma \) are not proportional in \( V \). This implies that

\[
\gamma - \beta \in A,
\]

as in Section 19.9.

If \( \gamma - \beta \in -A^+ \), so that \( \beta - \gamma \in A^+ \), then \( \beta = \gamma + (\beta - \gamma) \) would be the sum of two elements of \( A^+ \). This is not possible, because \( \beta \in B \). Otherwise, \( \gamma - \beta \in A^+ \), and of course

\[
\tau(\gamma - \beta) = \tau(\gamma) - \tau(\beta) = r - 1.
\]

This means that the induction hypothesis can be applied to \( \gamma - \beta \). Thus the desired statement for \( \gamma \) can be obtained from the analogous statement for \( \gamma - \beta \), with \( \alpha_r = \beta \).

Suppose now that \( A \) is reduced, as in Section 19.3. Let \( \alpha \in B \) be given, and let \( \sigma_\alpha \) be the symmetry on \( V \) with vector \( \alpha \) that maps \( A \) onto itself. Under these conditions,

\[
\sigma_\alpha(A^+ \setminus \{\alpha\}) = A^+ \setminus \{\alpha\},
\]
as in Proposition 6 on p32 of [23], and Lemma B on p50 of [13]. To see this, it suffices to verify that
\[(19.12.9) \quad \sigma_\alpha(A^+ \setminus \{\alpha\}) \subseteq A^+ \setminus \{\alpha\},\]
because $\sigma_\alpha$ is its own inverse on $V$.

Let $\gamma \in A^+ \setminus \{\alpha\}$ be given. Thus $\gamma$ can be expressed as
\[(19.12.10) \quad \gamma = \sum_{\beta \in B} m_\beta \beta,\]
where $m_\beta$ is a nonnegative integer for each $\beta \in B$, because $\gamma \in A^+$. Observe that $\gamma$ is not proportional to $\alpha$ in this situation, because $\gamma \neq \alpha$ by hypothesis, and $A$ is reduced. This means that there is a $\beta \in B$ such that $\beta \neq \alpha$ and $m_\beta \neq 0$.

Remember that $\sigma_\alpha(\gamma)$ is equal to $\gamma$ minus a multiple of $\alpha$. If we express $\sigma_\alpha(\gamma)$ as a linear combination of the basis vectors in $B$, we get that the coefficient of $\beta$ in the expression for $\sigma_\alpha(\gamma)$ is equal to $m_\beta$ too. This implies that $\sigma_\alpha(\gamma) \in A^+$, because $\sigma_\alpha(\gamma) \in A, m_\beta > 0$, and $B$ is a base for $A$. We also have that $\sigma_\alpha(\gamma) \neq \alpha$, because $m_\beta \neq 0$, as desired.

Put
\[(19.12.11) \quad \rho = \frac{1}{2} \sum_{\gamma \in A^+} \gamma,\]
which is an element of $V$. Let us check that
\[(19.12.12) \quad \sigma_\alpha(\rho) = \rho - \alpha,\]
as in the corollary on p32 of [23], and the corollary to Lemma B on p50 of [13].

Put
\[(19.12.13) \quad \rho_\alpha = \frac{1}{2} \sum_{\gamma \in A^+ \setminus \{\alpha\}} \gamma,\]
which is interpreted as being 0 when $A^+ = \{\alpha\}$. Using (19.12.8), we get that $\sigma_\alpha(\rho_\alpha) = \rho_\alpha$. Note that $\rho = \rho_\alpha + \alpha/2$, by construction. We also have that $\sigma_\alpha(\alpha) = -\alpha$. Thus
\[(19.12.14) \quad \sigma_\alpha(\rho) = \sigma_\alpha(\rho_\alpha) + \sigma_\alpha(\alpha)/2 = \rho_\alpha - \alpha/2 = \rho - \alpha,\]
as desired.

### 19.13 Bases and duality

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $A$ be a root system in $V$. If $\alpha \in A$, then we let $\sigma_\alpha$ be the symmetry on $V$ with vector $\alpha$ that maps $A$ onto itself, as usual. Let $\lambda_\alpha \in V'$ be the corresponding linear functional on $V$, so that $\lambda_\alpha(\alpha) = 2$ and $\lambda_\alpha = 0$ on the hyperplane of vectors in $V$ fixed by $\sigma_\alpha$. Remember that
\[(19.13.1) \quad A' = \{\lambda_\alpha : \alpha \in A\}\]
is a root system in the dual space $V'$, as in Section 19.8.

If $\alpha \in A$ and $2\alpha \in A$, then

$$\sigma_2 = \sigma_{2\alpha},$$

(19.13.2)

because $\sigma_{2\alpha}$ may also be considered as a symmetry on $V$ with vector $2\alpha$. In particular, this means that $\lambda_\alpha$ and $\lambda_{2\alpha}$ have the same kernel in $V$. One can check that

$$\lambda_{2\alpha} = \lambda_\alpha / 2,$$

(19.13.3)

using the condition that $\lambda_{2\alpha}(2\alpha) = 2$.

Let $B$ be a base for $A$, and put

$$B_1 = \{\beta \in B : 2\beta \not\in A\},$$

(19.13.4)

$$B_2 = \{\beta \in B : 2\beta \in A\}.$$

(19.13.5)

Of course, $B_2 = \emptyset$ when $A$ is reduced. We would like to show that

$$\{\lambda_\beta : \beta \in B_1\} \cup \{\lambda_\beta / 2 : \beta \in B_2\}$$

is a base for $A'$. This corresponds to Proposition 7 on p32 and the remark on p33 of [23]. Note that (19.13.6) is contained in $A'$, by (19.13.3).

Let $(v, w)$ be an inner product on $V$ that is invariant under the Weyl group of $A$, as in Section 19.4. Using $(.,.)$, we get an isomorphism between $V$ and $V'$ in the usual way, as in Section 19.7. If $\alpha \in V$ and $\alpha 
eq 0$, then put

$$\alpha = 2\alpha (\alpha, \alpha)^{-1},$$

(19.13.7)

as before. Clearly

$$\hat{\alpha} = 2\alpha / (\alpha, \alpha) = \hat{\alpha} / 2$$

(19.13.8)

in this case.

If $\alpha \in A$, then we have seen that $\hat{\alpha}$ corresponds to $\lambda_\alpha$ under the isomorphism between $V$ and $V'$ associated to $(.,.)$. Thus

$$\hat{A} = \{\hat{\alpha} : \alpha \in A\}$$

(19.13.9)

corresponds to $A'$ under this isomorphism, as before. We also have that $\hat{A}$ is a root system in $V$, as in Section 19.8.

We would like to show that

$$\{\hat{\beta} : \beta \in B_1\} \cup \{\hat{\beta} / 2 : \beta \in B_2\}$$

is a base for $\hat{A}$, as in [23]. This corresponds to Exercise 1 on p54 of [13], in the reduced case. Observe that (19.13.10) is contained in $\hat{A}$, because of (19.13.8). It will follow that (19.13.6) is a base for $A'$, because of the isomorphism between $V$ and $V'$ associated to $(.,.)$. It is easy to see that (19.13.10) is a basis for $V$, because $B$ is a basis for $V$.  


Let \( A^+ \) be the set of \( \alpha \in A \) that can be expressed as a linear combination of elements of \( B \) with coefficients that are nonnegative integers, as before. Also let \( \tau \) be a linear functional on \( V \) such that \( \tau(\beta) > 0 \) for every \( \beta \in B \). If \( \alpha \in A^+ \), then \( \tau(\alpha) > 0 \), and hence

\[
\tau(\tilde{\alpha}) > 0.
\]

Similarly, if \( \alpha \in -A^+ \), then \( \tau(\alpha) < 0 \), so that

\[
\tau(\tilde{\alpha}) < 0.
\]

In particular, \( \tau \) is nonzero at every element of \( \tilde{A} \).

Let \( \tilde{A}^+ \) be the set of elements of \( \tilde{A} \) on which \( \tau \) is positive, as in Section 19.11. Equivalently,

\[
(\tilde{A})^+_\tau = \{ \tilde{\alpha} : \alpha \in A^+ \},
\]

as in the preceding paragraph. Let us say that an element of \( (\tilde{A})^+_\tau \) is decomposable if it can be expressed as the sum of two elements of \( (\tilde{A})^+_\tau \), and indecomposable otherwise, as in Section 19.11. The set of indecomposable elements of \( (\tilde{A})^+_\tau \) is a base for \( \tilde{A} \), as before. Note that (19.13.10) is contained in \( (\tilde{A})^+_\tau \).

It suffices to show that (19.13.10) is the same as the set of indecomposable elements of \( (\tilde{A})^+_\tau \). If an element of (19.13.10) is decomposable as an element of \( (\tilde{A})^+_\tau \), then it could be expressed as the sum of two proportional elements of \( (\tilde{A})^+_\tau \). One can check that this is not possible, because of the way that \( B_1 \) and \( B_2 \) are defined. Thus (19.13.10) is contained in the set of indecomposable elements of \( (\tilde{A})^+_\tau \). It follows that these two sets are the same, because they have the same number of elements, which is the dimension of \( V \).

The argument in [23] considers the associated convex cones in \( V \). More precisely, let \( C \) be the set of vectors in \( V \) that can be expressed as linear combinations of elements of \( B \) with coefficients that are nonnegative real numbers. This is the same as the closed convex cone in \( V \) generated by \( A^+ \). This is also the same as the closed convex cone in \( V \) generated by \( (\tilde{A})^+_\tau \), by (19.13.13). The extremal rays in \( C \) correspond exactly to half-lines through elements of \( B \). Similarly, the extremal rays in \( C \) correspond to half-lines through the indecomposable elements of \( (\tilde{A})^+_\tau \). This means that these collections of half-lines in \( V \) are the same. It follows that the indecomposable elements of \( (\tilde{A})^+_\tau \) are all proportional to elements of \( B \). One can use this to check that the set of indecomposable elements of \( (\tilde{A})^+_\tau \) is the same as (19.13.10), as desired.

### 19.14 Bases and the Weyl group

Let \( V \) be a vector space over the real numbers of positive finite dimension, and let \( A \) be a reduced root system in \( V \). Also let \( \sigma_\alpha \) be the symmetry on \( V \) with vector \( \alpha \in A \) that maps \( A \) onto itself, as usual. Thus the Weyl group \( W \) of \( A \) is the subgroup of \( GL(V) \) generated by \( \sigma_\alpha \), \( \alpha \in A \). Let \( B \) be a base for \( A \), and let \( W_B \) be the subgroup of \( W \) generated by \( \sigma_\beta \), \( \beta \in B \). It is well known that

\[
W_B = W,
\]

(19.14.1)
as in Theorem 2 on p33 of [23], and the theorem on p51 of [13].

To show this, we first consider some other properties of \( W_B \), as in [13, 23].

Let \( \tau \) be a linear functional on \( V \). If \( \sigma \in W \), then let \( \sigma' \) be the corresponding dual linear mapping on the dual space \( V' \) of \( V \), so that

\[
\sigma'(\tau) = \tau \circ \sigma.
\]  

(19.14.2)

We would like to find a \( \sigma \in W_B \) such that

\[
(\sigma'(\tau))(\beta) = \tau(\sigma(\beta)) \geq 0
\]  

(19.14.3)

for every \( \beta \in B \). This corresponds to part (a) of Theorem 2 on p33 of [23], as mentioned in its proof. Part (a) of the theorem on p51 of [13] is very similar, but formulated a bit differently. See also Exercise 14 on p55 of [13].

Let \( A^+ \) be the set of \( \alpha \in A \) that can be expressed as a linear combination of elements of \( B \) with coefficients that are nonnegative integers, as usual. Also let \( \rho \) be the element of \( V \) that is one-half the sum of the elements of \( A^+ \). We can choose \( \sigma \in W_B \) such that

\[
(\sigma'(\tau))(\rho) = \tau(\sigma(\rho))
\]  

(19.14.4)

is maximal in \( R \). If \( \beta \in B \), then \( \sigma \circ \sigma_\beta \in W_B \), so that

\[
(\sigma'(\tau))(\rho) \geq ((\sigma \circ \sigma_\beta)'(\tau))(\rho).
\]  

(19.14.5)

Of course,

\[
((\sigma \circ \sigma_\beta)'(\tau))(\rho) = \tau(\sigma(\sigma_\beta(\rho))).
\]  

(19.14.6)

Under these conditions, we have that

\[
\sigma_\beta(\rho) = \rho - \beta,
\]  

(19.14.7)

as in (19.12.12). Thus

\[
((\sigma \circ \sigma_\beta)'(\tau))(\rho) = \tau(\sigma(\rho - \beta)) = (\sigma'(\tau))(\rho - \beta).
\]  

(19.14.8)

It follows that

\[
(\sigma'(\tau))(\rho) \geq (\sigma'(\tau))(\rho - \beta),
\]  

(19.14.9)

by (19.14.5). This implies (19.14.3), as desired.

If \( \tilde{B} \) is another base for \( A \), then there is a \( \sigma \in W_B \) such that

\[
\sigma(B) = \tilde{B}
\]  

(19.14.10)

This corresponds to part (b) of Theorem 2 on p33 of [23], and part (b) of the theorem on p51 of [13], as mentioned in their proofs again. To see this, let \( \tilde{\tau} \) be a linear functional on \( V \) such that

\[
\tilde{\tau}(\tilde{\beta}) > 0
\]  

(19.14.11)
for every $\tilde{\beta} \in \tilde{B}$, which is possible because $\tilde{B}$ is a basis for $V$. As in the previous paragraphs, there is a $\sigma \in W_B$ such that

$$\tag{19.14.12} (\sigma'(\tilde{\tau}))(\beta) \geq 0$$

for every $\beta \in B$. Put

$$\tag{19.14.13} \tau = \sigma'(\tilde{\tau}) = \tilde{\tau} \circ \sigma \in V',$$

so that

$$\tag{19.14.14} \tau(\beta) \geq 0$$

for every $\beta \in B$.

If $\alpha \in A$ can be expressed as a linear combination of elements of $\tilde{B}$ with coefficients that are nonnegative integers, then $\tilde{\tau}(\alpha) > 0$, as usual. Otherwise, $\tilde{\alpha} < 0$, so that $\tilde{\tau}(\alpha) \neq 0$ for every $\alpha \in A$. This implies that

$$\tag{19.14.15} \tau(\alpha) = \tilde{\tau}(\sigma(\alpha)) \neq 0$$

for every $\alpha \in A$, because $\sigma(A) = A$. It follows that $\tau(\beta) > 0$ for every $\beta \in B$, because of (19.14.14).

Let $A^+_\tau$ and $A^\pm_\tau$ be the collections of $\alpha \in A$ such that

$$\tag{19.14.16} \tilde{\tau}(\sigma(\alpha)) = \tau(\alpha) > 0$$

and $\tilde{\tau}(\alpha) > 0$, respectively. It is easy to see that

$$\tag{19.14.17} \sigma(A^+_\tau) = A^+_\tau,$$

because $\sigma(A) = A$. The elements of $B$ and $\tilde{B}$ may be characterized as the indecomposable elements of $A^+_\tau$ and $A^\pm_\tau$, respectively, as in Section 19.11. Thus (19.14.10) follows from (19.14.17), as desired.

If $\gamma \in A$, then there is a $\sigma \in W_B$ such that

$$\tag{19.14.18} \sigma(\gamma) \in B.$$

This corresponds to part (c) of Theorem 2 on p33 of [23], and part (c) of the theorem on p51 of [13], as mentioned in their proofs. In this argument, it will be convenient to let

$$\tag{19.14.19} V'_a = \{ \lambda \in V' : \lambda(\alpha) = 0 \}$$

be the hyperplane in $V'$ dual to a nonzero element $a$ of $V$. If $b$ is another nonzero element of $V$, then $V'_a = V'_b$ exactly when $a$ and $b$ are proportional in $V$. In particular, if $\alpha \in A$ and $\alpha \neq \pm \gamma$, then

$$\tag{19.14.20} V'_a \neq V'_b,$$

because $A$ is reduced, by hypothesis.

Thus, if $\alpha \in A$ and $\alpha \neq \pm \gamma$, then

$$\tag{19.14.21} V'_a \cap V'_\gamma.$$
19.15. Weyl chambers

Let \( V \) be a vector space over the real numbers of positive finite dimension \( n \), and let \( A \) be a root system in \( V \). Let us say that a linear functional \( \tau \) on \( V \) is regular with respect to \( A \) if for every \( \alpha \in A \), \( \tau(\alpha) \neq 0 \). This corresponds to terminology on p48 of [13], although in [13] an inner product on \( V \) is used, so that the dual space \( V' \) of \( V \) can be identified with \( V \). The set of linear functionals on \( V \) that are regular with respect to \( A \) is the same as

\[
V' \setminus \left( \bigcup_{\alpha \in A} V'_{\alpha} \right),
\]

19.15. Weyl chambers

Let \( V \) be a vector space over the real numbers of positive finite dimension \( n \), and let \( A \) be a root system in \( V \). Let us say that a linear functional \( \tau \) on \( V \) is regular with respect to \( A \) if for every \( \alpha \in A \), \( \tau(\alpha) \neq 0 \). This corresponds to terminology on p48 of [13], although in [13] an inner product on \( V \) is used, so that the dual space \( V' \) of \( V \) can be identified with \( V \). The set of linear functionals on \( V \) that are regular with respect to \( A \) is the same as

\[
V' \setminus \left( \bigcup_{\alpha \in A} V'_{\alpha} \right),
\]
where $V'$ is the hyperplane in $V'$ dual to $\alpha \in A$, as in (19.14.19). Of course, $V'$ is isomorphic to $\mathbb{R}^n$, as a vector space over $\mathbb{R}$. This leads to a natural topology on $V'$, corresponding to the standard topology on $\mathbb{R}^n$. This topology does not depend on the particular isomorphism with $\mathbb{R}^n$, because linear mappings on $\mathbb{R}^n$ are continuous with respect to the standard topology. Note that (19.15.1) is an open set with respect to this topology. Remember that (19.15.1) is nonempty, as in Section 18.12.

Let $B$ be a base for $A$, and let $A^{B,+}$ be the set of $\alpha \in A$ that can be expressed as a linear combination of elements of $B$ with coefficients that are nonnegative integers. Thus

$$A = A^{B,+} \cup (-A^{B,+}),$$

by the definition of a base. If $\tau \in V'$ satisfies

$$\tau(\beta) > 0 \quad \text{for every } \beta \in B,$$

then

$$\tau(\alpha) > 0 \quad \text{for every } \alpha \in A^{B,+},$$

as before. The converse holds trivially, because $B \subseteq A^{B,+}$. Observe that $\tau$ is regular with respect to $A$ when (19.15.4) holds, because of (19.15.2). The set of $\tau \in V'$ that satisfy (19.15.3) is called the Weyl chamber associated to $B$, as in Remark 2 on p34 of [23]. This corresponds to the fundamental Weyl chamber relative to $B$ defined on p49 of [13], using an inner product on $V$, as before.

Let $\tau$ be a linear functional on $V$ that is regular with respect to $A$, and let $A^+_B$ be the set of $\alpha \in A$ such that $\tau(\alpha) > 0$, as in Section 19.11. If $\tau$ satisfies (19.15.3) and hence (19.15.4), then $A^{B,+} \subseteq A^+_B$, which implies that

$$A^{B,+} = A^+_B,$$

because of (19.15.2), as in Section 19.11. Conversely, (19.15.5) implies (19.15.4), and thus (19.15.3). Let $B_\tau$ be the set of indecomposable elements of $A^+_B$, which is a base for $A$, as in Section 19.11. If $\tau$ satisfies (19.15.3), then $B = B_\tau$, as before. Conversely, if $B = B_\tau$, then $B \subseteq A^+_B$, which means that (19.15.3) holds. This shows that the Weyl chamber associated to $B$ consists exactly of the linear functionals $\tau$ on $V$ that are regular with respect to $A$ and satisfy $B = B_\tau$. In particular, if $\tau$ is any linear functional on $V$ that is regular with respect to $A$, then $\tau$ is an element of the Weyl chamber associated to $B_\tau$.

It is easy to see that the Weyl chamber associated to $B$ is an open convex cone in $V$, so that it is connected in particular. If $\tilde{B}$ is another base for $A$, $\tilde{B} \neq B$, then the Weyl chambers associated to $B$ and $\tilde{B}$ are disjoint, by the remarks in the preceding paragraph. The union of the Weyl chambers associated to bases for $A$ is equal to the set of linear functionals on $V$ that are regular with respect to $A$, as before. It follows that the Weyl chambers associated to bases for $A$ are the same as the connected components of (19.15.1).

If $\tilde{B}$ is another base for $A$ and

$$\tilde{B} \subseteq A^{B,+},$$

(19.15.6)
then
\[ (19.15.7) \quad \tilde{B} = B. \]
Indeed, if \( \tau \in V' \) satisfies (19.15.3) and hence (19.15.4), then \( \tau(\tilde{\beta}) > 0 \) for every \( \tilde{\beta} \in \tilde{B} \). This implies that \( \tilde{B} = B_\tau \), as before, so that (19.15.7) holds, because \( B = B_\tau \) too.
Chapter 20

Root systems, 2

20.1 Products in the Weyl group

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $A$ be a reduced root system in $V$. If $\alpha \in A$, then we let $\sigma_\alpha$ be the symmetry on $V$ with vector $\alpha$ that maps $A$ onto itself, as before. Let $B$ be a base for $A$, and remember that the Weyl group $W$ of $A$ is generated by $\sigma_\beta$, $\beta \in B$, as in Section 19.14.

Let $r \geq 2$ be an integer, and let $\gamma_1, \ldots, \gamma_r$ be $r$ elements of $B$, possibly with repetitions. Put $\sigma_j = \sigma_{\gamma_j}$ for each $j = 1, \ldots, r$, and consider

\[(\sigma_1 \circ \cdots \circ \sigma_1) (\gamma_r),\]

which is an element of $A$. Thus (20.1.1) can be expressed as a linear combination of elements of $B$ with integer coefficients, and where the coefficients are either all greater than or equal to 0, or all less than or equal to 0. If the coefficients are all less than or equal to 0, then there is a positive integer $n < r$ such that

\[(\sigma_1 \circ \cdots \circ \sigma_1) (\gamma_r) = (\sigma_1 \circ \cdots \circ \sigma_n \circ \sigma_{n+1} \circ \cdots \circ \sigma_{r-1}) (\gamma_r),\]

This corresponds to Lemma C on p50 of [13]. More precisely, in the first part of the right side of (20.1.2), $\sigma_1 \circ \cdots \circ \sigma_{n-1}$ is interpreted as being the identity mapping on $V$ when $n = 1$. Similarly, if $n = r - 1$, then $\sigma_{n+1} \circ \cdots \circ \sigma_{r-1}$ is interpreted as being the identity mapping on $V$ in the second part of the right side of (20.1.2).

To see this, put

\[(\sigma_1 \circ \cdots \circ \sigma_1) (\gamma_r) = (\sigma_{j+1} \circ \cdots \circ \sigma_{r-1}) (\gamma_r)\]

for $j = 0, \ldots, r - 2$, and $\delta_{r-1} = \gamma_r$. Note that $\delta_0$ is the same as (20.1.1), whose coefficients with respect to $B$ are all less than or equal to 0, by hypothesis. Of course, the coefficients of $\delta_{r-1} = \gamma_r$ with respect to $B$ are all greater than or equal to 0, because $\gamma_r \in B$. Let $n$ be the smallest positive integer less than or
equal to 0 such that the coefficients of $\delta_n$ with respect to $B$ are all greater than or equal to 0, so that the coefficients of
\begin{equation}
\delta_{n-1} = \sigma_n(\delta_n)
\end{equation}
with respect to $B$ are all less than or equal to 0. This implies that
\begin{equation}
\gamma_n = \delta_n,
\end{equation}
by (19.12.8).

If $\alpha \in A$ and $T$ is an invertible linear mapping on $V$ that maps $A$ onto itself, then $T \circ \sigma_{\alpha} \circ T^{-1} = \sigma_{T(\alpha)}$, as in Section 19.4. Let us take $\alpha = \gamma_r$ and $T = \sigma_{n+1} \circ \cdots \circ \sigma_{r-1}$, so that $T(\alpha) = \delta_n$. Combining this with (20.1.5), we obtain that
\begin{equation}
\sigma_n = (\sigma_{n+1} \circ \cdots \circ \sigma_{r-1}) \circ \sigma_r \circ (\sigma_{r-1} \circ \cdots \circ \sigma_{n+1}).
\end{equation}
It is easy to get (20.1.2) from (20.1.6).

Put
\begin{equation}
\sigma = \sigma_1 \circ \cdots \circ \sigma_r,
\end{equation}
and suppose now that $r$ is the smallest positive integer such that $\sigma$ can be expressed as the composition of $r$ symmetries associated to elements of $B$. The previous statement implies that the coefficients of (20.1.1) with respect to $B$ are all greater than or equal to 0. Remember that $\sigma_r(\gamma_r) = -\gamma_r$, so that
\begin{equation}
\sigma(\gamma_r) = -(\sigma_1 \circ \cdots \circ \sigma_{r-1})(\gamma_r).
\end{equation}
It follows that the coefficients of $\sigma(\gamma_r)$ with respect to $B$ are all less than or equal to 0, as in the corollary to Lemma C on p50 of [13]. Note that this holds automatically when $r = 1$.

If $\sigma$ is any element of the Weyl group $W$ of $A$, then $\sigma$ can be expressed as the composition of finitely many symmetries associated to elements of $B$, as before. In particular, if $\sigma$ is not the identity mapping on $V$, then one needs at least one of these symmetries. This means that $\sigma$ can be expressed as in (20.1.7) for some positive integer $r$, and we may as well take $r$ to be as small as possible. Under these conditions, $\sigma$ cannot map $B$ onto itself, as in the preceding paragraph.

Equivalently, if $\sigma \in W$ satisfies
\begin{equation}
\sigma(B) = B,
\end{equation}
then $\sigma$ is the identity mapping on $V$. This corresponds to part (e) of the theorem on p51 of [13]. This also corresponds to Proposition 4 on p62 of [23].

## 20.2 The Cartan matrix

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $A$ be a root system in $V$. If $\alpha \in A$, then let $\sigma_\alpha$ be the symmetry on $V$
with vector $\alpha$ that maps $A$ onto itself, and let $\lambda_\alpha$ be the corresponding linear functional on $V$. Thus $\lambda_\alpha(\alpha) = 2$, and $\lambda_\alpha = 0$ on the hyperplane of vectors in $V$ that are fixed by $\sigma_\alpha$. If $\beta \in A$ too, then we put

$$n(\alpha, \beta) = \lambda_\beta(\alpha),$$

as in Section 19.9. Of course, this is an integer, by the definition of a root system.

Let $(\cdot, \cdot)$ be an inner product on $V$ that is invariant under the Weyl group $W$ of $A$. This means that $\sigma_\alpha$ is the reflection on $V$ with respect to $(\cdot, \cdot)$ associated to $\alpha \in A$, so that $\lambda_\alpha(v) = 2(\langle \alpha, \alpha \rangle^{-1} v, \alpha)$ for every $v \in V$. In particular,

$$n(\alpha, \beta) = 2(\langle \alpha, \beta \rangle, \beta)^{-1}$$

for every $\alpha, \beta \in A$.

Let $B$ be a base for $A$. The Cartan matrix of $A$ with respect to $B$ is $n(\alpha, \beta)$ as a function of $\alpha, \beta \in B$, as in Definition 5 on p34 of [23], and p55 of [13]. Note that $n(\alpha, \alpha) = \lambda_\alpha(\alpha) = 2$ for every $\alpha \in A$, and in particular for $\alpha \in B$. If $\alpha, \beta \in B$ and $\alpha \neq \beta$, then $n(\alpha, \beta) \leq 0$, by (19.11.8) and (20.2.2). Of course, $\alpha$ and $\beta$ are not proportional in this situation. It follows that

$$n(\alpha, \beta) \in \{0, -1, -2, -3\}$$

when $\alpha \neq \beta$, as in Section 19.9, and mentioned on p34 of [23]. If $A$ is reduced, then $A$ is determined up to isomorphism by the Cartan matrix, as in Proposition 8 on p34 of [23].

More precisely, let $\tilde{V}$ be another vector space over the real numbers of positive finite dimension, and let $\tilde{A}$ be a reduced root system in $\tilde{V}$. If $\tilde{\alpha} \in \tilde{A}$, then let $\tilde{\sigma}_\alpha$ on $\tilde{V}$ with vector $\tilde{\alpha}$ that maps $\tilde{A}$ onto itself, as usual. Also let $\tilde{\lambda}_\alpha$ be the corresponding linear functional on $\tilde{V}$, so that $\tilde{\lambda}_\alpha(\tilde{\alpha}) = 2$ and $\tilde{\lambda}_\alpha = 0$ on the hyperplane of vectors in $\tilde{V}$ that are fixed by $\tilde{\sigma}_\alpha$. If $\tilde{\beta} \in \tilde{A}$ too, then we put

$$\tilde{n}(\tilde{\alpha}, \tilde{\beta}) = \tilde{\lambda}_\beta(\tilde{\alpha}),$$

as before. Let $\tilde{B}$ be a base for $\tilde{A}$, so that the restriction of (20.2.4) to $\tilde{\alpha}, \tilde{\beta} \in \tilde{B}$ is the Cartan matrix of $\tilde{A}$ with respect to $\tilde{B}$.

Suppose that $\phi$ is a one-to-one mapping from $B$ onto $\tilde{B}$ such that

$$\tilde{n}(\phi(\alpha), \phi(\beta)) = n(\alpha, \beta)$$

for every $\alpha, \beta \in B$. If $A$ is reduced, then $\phi$ extends to a unique one-to-one linear mapping $f$ from $V$ onto $\tilde{V}$ such that

$$f(A) = \tilde{A}.$$
vector spaces over $\mathbb{R}$, by the definition of a base for a root system. Thus $\phi$ automatically extends to a unique one-to-one linear mapping $f$ from $V$ onto $\tilde{V}$, and in particular $V$ and $\tilde{V}$ have the same dimension as vector spaces over $\mathbb{R}$.

If $\alpha, \beta \in B$, then

$$(20.2.7) \quad (\tilde{\sigma}_{\phi(\alpha)} \circ f)(\beta) = \tilde{\sigma}_{\phi(\alpha)}(\phi(\beta)) = \phi(\beta) - \tilde{n}(\phi(\beta), \phi(\alpha)) \phi(\alpha)$$

and

$$(20.2.8) \quad (f \circ \sigma_{\alpha})(\beta) = f(\beta - n(\beta, \alpha) \alpha) = \phi(\beta) - n(\beta, \alpha) \phi(\alpha).$$

Using our hypothesis (20.2.5), we get that the right sides of (20.2.7) and (20.2.8) are the same. This implies that

$$(20.2.9) \quad \tilde{\sigma}_{\phi(\alpha)} \circ f = f \circ \sigma_{\alpha}$$

on $V$ for every $\alpha \in B$, because $V$ is spanned by $\beta \in B$. Equivalently,

$$(20.2.10) \quad \tilde{\sigma}_{\phi(\alpha)} = f \circ \sigma_{\alpha} \circ f^{-1}$$

on $\tilde{V}$ for every $\alpha \in B$.

As usual, $GL(V)$ and $GL(\tilde{V})$ denote the general linear groups of invertible linear mappings on $V$ and $\tilde{V}$, respectively. Of course,

$$(20.2.11) \quad \sigma \mapsto f \circ \sigma \circ f^{-1}$$

is a group isomorphism from $GL(V)$ onto $GL(\tilde{V})$. Remember that the Weyl groups $W$ and $\tilde{W}$ of $A$ and $\tilde{A}$ are generated by the symmetries associated to elements of $B$ and $\tilde{B}$, respectively, as in Section 19.14. Using this and (20.2.10), we get that (20.2.11) maps $W$ onto $\tilde{W}$.

Remember that $A$ is the same as the set of images of elements of $B$ under elements of $W$, and similarly for $\tilde{A}$, as in Section 19.14. Using this and the remarks in the preceding paragraph, it is easy to see that (20.2.6) holds, as desired.

Of course, a one-to-one linear mapping $f$ from $V$ onto $\tilde{V}$ that satisfies (20.2.6) is the same as an isomorphism between the root systems $A$ and $\tilde{A}$, as in Section 19.5. In this case, we have that

$$(20.2.12) \quad \tilde{n}(f(\alpha), f(\beta)) = n(\alpha, \beta)$$

for every $\alpha, \beta \in A$, by (19.5.9).

Let $\text{Aut}(A)$ be the group of automorphisms of $A$, which is the subgroup of $GL(V)$ consisting of invertible linear mappings on $V$ that send $A$ onto itself. Remember that the Weyl group $W$ of $A$ is a normal subgroup of $\text{Aut}(A)$, as in Section 19.4. If $T \in \text{Aut}(A)$, then

$$(20.2.13) \quad n(T(\alpha), T(\beta)) = n(\alpha, \beta)$$

for every $\alpha, \beta \in A$, as in (20.2.12). Of course,

$$(20.2.14) \quad \{T \in \text{Aut}(A) : T(B) = B\}$$
is a subgroup of Aut(A) as well. The elements of this subgroup are uniquely determined by their restrictions to B, because B is a basis for V as a vector space over \( \mathbb{R} \). If A is reduced, then the restrictions of elements of (20.2.14) to B are characterized by the condition that (20.2.13) hold for every \( \alpha, \beta \in B \), as before. In this case, the intersection of W and (20.2.14) is the trivial subgroup of Aut(A), as in the previous section.

If \( T \) is any element of Aut(A), then it is easy to see that \( T(B) \) is a base for A too. If A is reduced, then there is a \( \sigma \in W \) such that

\[
\sigma(T(B)) = B,
\]

as in Section 19.14. Equivalently, this means that \( \sigma \circ T \) is an element of (20.2.14). This corresponds to Proposition 9 on p35 of [23], and some remarks beginning on p65 of [13].

### 20.3 Coxeter graphs

As on p35 of [23], a **Coxeter graph** is a finite graph, where each pair of distinct vertices may be joined by 0, 1, 2, or 3 edges.

Let V be a vector space over the real numbers of positive finite dimension, let A be a root system in V, and let B be a base for A. If \( \alpha, \beta \in A \), then \( \lambda_\alpha \in V' \) and \( n(\alpha, \beta) \in \mathbb{Z} \) are defined in the usual way, as in the previous section. If \( \alpha, \beta \in B \) and \( \alpha \neq \beta \), then \( \alpha \) and \( \beta \) are not proportional, and

\[
n(\alpha, \beta) = n(\beta, \alpha)
\]

can only be equal to 0, 1, 2, or 3, as in Section 19.9. Of course, (20.3.1) is symmetric in \( \alpha \) and \( \beta \).

The **Coxeter graph of A with respect to B** is defined as follows, as on p35 of [23], and p56 of [13]. We take B to be the set of vertices of the Coxeter graph. If \( \alpha \) and \( \beta \) are distinct elements of B, then the number of edges between \( \alpha \) and \( \beta \) is equal to (20.3.1).

If \( \tilde{B} \) is another base for A, then there is an element \( \sigma \) of the Weyl group of A that maps B onto \( \tilde{B} \), as in Section 19.14. This leads to an isomorphism between the Coxeter graphs of A with respect to B and \( \tilde{B} \).

Let \( \langle \cdot, \cdot \rangle \) be an inner product on V that is invariant under the Weyl group of A. Using (20.2.2), we get that

\[
n(\alpha, \beta) = 4(\alpha, \beta)^2 \|\alpha\|^{-2} \|\beta\|^{-2}
\]

for every \( \alpha, \beta \in A \), where \( \|\cdot\| \) is the norm on V associated to \( \langle \cdot, \cdot \rangle \). Equivalently, this is the same as 4 times the square of the cosine of the angle between \( \alpha \) and \( \beta \), as in Section 19.9. In particular, (20.3.2) is equal to 0 exactly when \( \langle \alpha, \beta \rangle = 0 \).

The Coxeter graph of A with respect to B is determined by the Cartan matrix of A with respect to B in an obvious way. Remember that the diagonal entries of the Cartan matrix are equal to 2, and that the other entries are less
than or equal to 0. It is easy to see that the Cartan matrix of $A$ with respect to $B$ is determined by the Coxeter graph of $A$ with respect to $B$ together with the ratios of the norms of the elements of $B$.

More precisely, let $\alpha$ and $\beta$ be distinct elements of $B$. If (20.3.2) is equal to 0, then $(\alpha, \beta) = 0$, and hence $n(\alpha, \beta) = 0$. Otherwise, if (20.3.2) is not zero, and if one knows if $\|\alpha\|$ and $\|\beta\|$ are the same, or which is larger, then $n(\alpha, \beta)$ can be determined from (20.3.2) as in Section 19.9. This is discussed beginning on p56 of [13], and on p38 of [23].

20.4 Reducibility

Let $V$ be a vector space over the real numbers of positive finite dimension again, let $A$ be a root system in $V$, and let $(\cdot, \cdot)$ be an inner product on $V$ that is invariant under the Weyl group of $A$. Suppose for the moment that $V$ is the direct sum of two nontrivial linear subspaces $V_1$ and $V_2$, as a vector space over $\mathbb{R}$, and that

\begin{equation}
A \subseteq V_1 \cup V_2.
\end{equation}

Put $A_j = A \cap V_j$ for $j = 1, 2$. If $\alpha \in V_1$ and $\beta \in V_2$, then $\alpha - \beta$ is not an element of $V_1 \cup V_2$, and thus is not in $A$. This implies that $(\alpha, \beta) \leq 0$, as in Section 19.9. The same argument could be applied to $\alpha$ and $-\beta$, to get that $(\alpha, \beta) \geq 0$, and hence

\begin{equation}
(\alpha, \beta) = 0.
\end{equation}

It is easy to see that the linear span of $A_j$ is equal to $V_j$ for each $j = 1, 2$, because the linear span of $A$ is $V$. It follows that the elements of $V_1$ and $V_2$ are orthogonal to each other with respect to $(\cdot, \cdot)$, as in part (a) of Proposition 10 on p36 of [23].

If $\alpha \in A$, then the symmetry $\sigma_{\alpha}$ on $V$ with vector $\alpha$ that maps $A$ onto itself is the reflection associated to $\alpha$ with respect to $(\cdot, \cdot)$. If $\alpha \in A_1$, then $\sigma_{\alpha}$ fixes every element of the hyperplane orthogonal to $\alpha$, which contains $V_2$. In particular, $\sigma_{\alpha}(V_2) = V_2$, which implies that $\sigma_{\alpha}(V_1) = V_1$. One also can check the latter more directly. It follows that $A_1$ is a root system in $V_1$, and similarly $A_2$ is a root system in $V_2$. This is part (b) of Proposition 10 on p36 of [23]. Under these conditions, $A$ is considered to be the sum of $A_1$ and $A_2$, as on p36 of [23]. If $A$ cannot be expressed as the sum of two root systems in this way, then $A$ is said to be irreducible, as on p36 of [23], and p52 of [13].

Irreducibility of $A$ is equivalent to the connectedness of the Coxeter graph of $A$, as in Proposition 12 on p36 of [23], and mentioned on p57 of [13]. More precisely, the connectedness of the Coxeter graph of $A$ with respect to a base $B$ for $A$ does not depend on the choice of $B$, because other choices lead to isomorphic graphs, as before.

Suppose that $A$ corresponds to the sum of subsystems $A_1$ and $A_2$, as before. If $B_1$ and $B_2$ are bases for $A_1$ and $A_2$, then it is easy to see that $B_1 \cup B_2$ is a base for $A$. If $\alpha \in B_1$ and $\beta \in B_2$, then $(\alpha, \beta) = 0$, so that $\alpha$ and $\beta$ are not connected by any edges in the Coxeter graph of $A$ with respect to $B_1 \cup B_2$. This
means that the Coxeter graph of $A$ with respect to $B_1 \cup B_2$ is not connected, as desired. This is one half of the proof of Proposition 12 on p36 of [23].

Alternatively, let $B$ be any base for $A$, and put $B_j = A_j \cap B$ for $j = 1, 2$. Every element of $B_1$ is orthogonal to every element of $B_2$, so that there are no edges between $B_1$ and $B_2$ in the Coxeter graph of $A$ with respect to $B$. To get that the Coxeter graph is not connected, one has to check that $B_1, B_2 \neq \emptyset$. Equivalently, one can verify that $B$ cannot be contained in $A_1$ or $A_2$, as on p52 of [13].

Conversely, suppose that the Coxeter graph of $A$ with respect to a base $B$ is not connected. This implies that $B$ can be expressed as the union of disjoint nonempty subsets $B_1$ and $B_2$, with no edges in the Coxeter graph between elements of $B_1$ and $B_2$. Equivalently, every element of $B_1$ is orthogonal to every element of $B_2$ with respect to $(\cdot, \cdot)$. Let $V_1$ and $V_2$ be the linear subspaces of $V$ spanned by $B_1$ and $B_2$, respectively, so that every element of $V_1$ is orthogonal to every element of $V_2$. Note that $V$ corresponds to the direct sum of $V_1$ and $V_2$, because $V$ is spanned by $B$.

If $\beta \in B_1$, then $\sigma_\beta$ fixes every element of the hyperplane orthogonal to $\beta$, which contains $V_2$. Thus $\sigma_\beta(V_2) = V_2$, which implies that $\sigma_\beta(V_1) = V_1$. Alternatively, one can check directly that $\sigma_\beta(V_1) = V_1$, which implies that $\sigma_\beta(V_2) = V_2$. Similarly, if $\beta \in B_2$, then $\sigma_\beta$ maps $V_1$ and $V_2$ onto themselves.

Remember that every element $\alpha$ of $A$ can be obtained from elements of $B$ by compositions of the reflections $\sigma_\beta$, $\beta \in B$, as in Section 19.14. It follows that $\alpha$ is contained in $V_1$ or $V_2$. This implies that $A$ is reducible, as on p37 of [23], and p52 of [13].

### 20.5 Root strings

Let $V$ be a vector space over the real numbers of positive finite dimension, let $A$ be a root system in $V$, and let $(\cdot, \cdot)$ be an inner product on $V$ that is invariant under the Weyl group of $A$. Suppose that $\alpha$ and $\beta$ are nonproportional elements of $A$. Consider the elements of $A$ of the form

$$\beta + j \alpha, \quad (20.5.1)$$

where $j$ is an integer. The collection of these elements of $A$ is called the $\alpha$-string through $\beta$ in $A$, as on p45 of [13]. Let $r$ and $q$ be the largest integers such that

$$\beta - r \alpha, \beta + q \alpha \in A, \quad (20.5.2)$$

respectively. Thus $r, q \geq 0$, because $\beta \in A$, by hypothesis. We would like to show that if $j \in \mathbb{Z}$ satisfies $-r < j < q$, then (20.5.1) is an element of $A$ too.

Otherwise, there are integers $j_1, j_2$ such that $-r \leq j_1 < j_2 \leq q$ such that

$$\beta + j_1 \alpha, \beta + j_2 \alpha \in A \quad (20.5.3)$$

and

$$\beta + (j_1 + 1) \alpha, \beta + (j_2 - 1) \alpha \notin A. \quad (20.5.4)$$
Observe that neither $\beta + j_1 \alpha$ nor $\beta + j_2 \alpha$ is proportional to $\alpha$, because $\beta$ is not proportional to $\alpha$, by hypothesis. It follows that

\[(\alpha, \beta + j_1 \alpha) \geq 0, \quad (\alpha, \beta + j_2 \alpha) \leq 0,\]

as in Section 19.9. However,

\[(j_2 - j_1) (\alpha, \alpha) > 0,\]

because $j_1 < j_2$ and $\alpha \neq 0$. This is a contradiction, as on p45 of [13].

The fact that (20.5.1) is an element of $A$ when $j \in \mathbb{Z}$ satisfies $-r < j < q$ also came up in the Lie algebra setting in Section 17.9, as in part (e) of the proposition on p39 of [13].

Let $\sigma_\alpha$ be the symmetry on $V$ with vector $\alpha$ that maps $A$ onto itself, which is the reflection on $V$ associated to $\alpha$ with respect to $(\cdot, \cdot)$. Remember that $\sigma_\alpha(\beta) - \beta$ is an integer multiple of $\alpha$, by the definition of a root system. If $j \in \mathbb{Z}$, then

\[\sigma_\alpha(\beta + j \alpha) = \sigma_\alpha(\beta) - j \alpha\]

is the sum of $\beta$ and an integer multiple of $\alpha$ as well. Thus $\sigma_\alpha$ maps the $\alpha$-string through $\beta$ onto itself, because $\sigma_\alpha(A) = A$. Using this, it is easy to see that

\[\sigma_\alpha(\beta + q \alpha) = \beta - r \alpha.\]

Equivalently, $\sigma_\alpha(\beta) - q \alpha = \beta - r \alpha$, so that

\[\sigma_\alpha(\beta) - \beta = (q - r) \alpha.\]

This means that

\[2 (\beta, \alpha) (\alpha, \alpha)^{-1} = r - q,\]

by the usual expression for $\sigma_\alpha$ in terms of $\alpha$ and $(\cdot, \cdot)$. It follows that

\[|r - q| \leq 3,\]

because $\alpha$ and $\beta$ are not proportional, as in Section 19.9.

Put $\beta_0 = \beta - r \alpha$, which is an element of $A$ that is not proportional to $\alpha$. Observe that the $\alpha$-string through $\beta_0$ consists of the same elements of $A$ as the $\alpha$-string through $\beta$. Let $r_0$, $q_0$ be the largest integers such that $\beta_0 - r_0 \alpha$ and $\beta_0 + q_0 \alpha$ are elements of $A$, as before. By construction, $r_0 = 0$ and $q_0 = q + r$. We also have that $|r_0 - q_0| \leq 3$, as in (20.5.11). This means that

\[q + r \leq 3.\]

Thus the $\alpha$-string through $\beta$ has at most 4 elements, as on p45 of [13].
20.6 Roots and linear subspaces

Let $V$ be a vector space over the real numbers of positive finite dimension, as usual. Suppose for the moment that $\alpha$ is a nonzero element of $V$, and that $\sigma_{\alpha}$ is a symmetry on $V$ with vector $\alpha$. If $V_0$ is a linear subspace of $V$, then we might like to know if

\begin{equation}
\sigma_{\alpha}(V_0) \subseteq V_0.
\end{equation}

Equivalently, this means that

\begin{equation}
\sigma_{\alpha}(V_0) = V_0;
\end{equation}

because $\sigma_{\alpha}$ is its own inverse on $V$. If $\alpha \in V_0$, then it is easy to see that (20.6.1) holds, and hence (20.6.2) holds as well. Of course, (20.6.2) holds when $V_0$ is contained in the hyperplane of vectors in $V$ that are fixed by $\sigma_{\alpha}$. In fact, these are the only two ways in which (20.6.2) can hold, as in Exercise 1 on p45 of [13]. Indeed, if (20.6.1) holds, and if there is a $v_0 \in V_0$ such that $\sigma_{\alpha}(v_0) \neq v_0$, then $\sigma_{\alpha}(v_0) - v_0$ is a nonzero multiple of $\alpha$ that is contained in $V_0$.

Let $A$ be a root system in $V$, and for each $\alpha \in A$, let $\sigma_{\alpha}$ be the symmetry on $V$ with vector $\alpha$ that maps $A$ onto itself. Let $V_0$ be a nontrivial linear subspace of $V$, and put

\begin{equation}
A_0 = A \cap V_0.
\end{equation}

If $V_0$ is spanned by $A_0$, then $A_0$ is a root system in $V_0$. This follows from the fact that (20.6.2) holds for every $\alpha \in A_0$. This is related to Exercise 7 on p46 of [13].

Let $A_1$ be a nonempty subset of $A$ that satisfies the following two conditions. First, if $\alpha \in A_1$, then $-\alpha \in A_1$. Second, if $\alpha, \beta \in A_1$ and $\alpha + \beta \in A$, then

\begin{equation}
\alpha + \beta \in A_1.
\end{equation}

Under these conditions, $A_1$ is a root system in the linear subspace $V_1$ of $V$ that it spans, as in Exercise 7 on p46 of [13]. Note that this includes the situation mentioned in the preceding paragraph. One way to see this is to check that if $\alpha$ and $\beta$ are nonproportional elements of $A_1$, then $A_1$ contains the $\alpha$-string through $\beta$ in $A$, using the remarks in the previous section. This implies that $\sigma_{\alpha}(\beta) \in A_1$ in this case. Of course, if $\alpha, \beta \in A_1$ are proportional, then $\sigma_{\alpha}(\beta) = -\beta \in A_1$, by the first condition. It follows that for every $\alpha \in A_1$, $\sigma_{\alpha}$ maps $A_1$ into itself, and hence $\sigma_{\alpha}(A_1) = A_1$.

Let $(\cdot, \cdot)$ be an inner product on $V$ that is invariant under the Weyl group of $A$. If $c$ is a positive real number, then let $A(c)$ be the set of $\alpha \in A$ such that $(\alpha, \alpha) = c$. If $A(c) \neq \emptyset$, then one can verify that $A(c)$ is a root system in the linear subspace $V(c)$ of $V$ that it spans. This is part of Exercise 11 on p47 of [13].

20.7 More on irreducibility

Let $V$ be a vector space over the real numbers of positive finite dimension, let $A$ be an irreducible root system in $V$, and let $(\cdot, \cdot)$ be an inner product on $V$.
that is invariant under the Weyl group $W$ of $A$. We would like to check that $W$ acts irreducibly on $V$, as in the first part of Lemma B on p53 of [13]. More precisely, this means that there is no nontrivial proper linear subspace of $V$ that is mapped into itself by the elements of $W$. To see this, let $V_1$ be a nonzero linear subspace of $V$ that is mapped into itself by elements of $W$. If $\alpha \in A$, then let $\sigma_\alpha$ be the reflection on $V$ associated to $\alpha$ with respect to $(\cdot, \cdot)$, as usual. By hypothesis,

$$\sigma_\alpha(V_1) \subseteq V_1,$$

so that either $\alpha \in V_1$ or $V_1$ is contained in the hyperplane of vectors that are fixed by $\sigma_\alpha$, as in the previous section. Remember that the hyperplane of vectors that are fixed by $\sigma_\alpha$ is orthogonal to $\alpha$. In the second case, we get that $\alpha$ is contained in the orthogonal complement $V_2$ of $V_1$ in $V$. Thus $A \subseteq V_1 \cup V_2$, which implies that $V_2 = \{0\}$, because $A$ is supposed to be irreducible in $V$. This means that $V_1 = V$, as desired.

If $v \in V$, then the linear span of the set of images of $v$ under elements of $W$ is a linear subspace of $V$ that is mapped into itself by elements of $W$. If $v \neq 0$, then it follows that this linear span is equal to $V$, as in the second part of Lemma B on p53 of [13]. If $u$ is another nonzero element of $V$, then we get that

$$\langle \sigma(v), u \rangle \neq 0$$

for some $\sigma \in W$.

Suppose that $\alpha, \beta \in A$ satisfy

$$\|\alpha\| = \|\beta\|,$$

where $\|\cdot\|$ is the norm associated to $(\cdot, \cdot)$ on $V$. We would like to show that $\alpha$ is the image of $\beta$ under an element of $W$, as in the second part of Lemma C on p53 of [13]. We may as well suppose that $(\alpha, \beta) \neq 0$, by replacing $\beta$ by one of its images under an element of $W$ if necessary, as in the preceding paragraph. If $\alpha$ and $\beta$ are proportional, then $\alpha = \pm \beta$, by (20.7.3). There is nothing to do when $\alpha = \beta$, and if $\alpha = -\beta$, then $\alpha = \sigma_\beta(\beta)$, as desired. Suppose that $\alpha$ and $\beta$ are not proportional, and let $n(\alpha, \beta), n(\beta, \alpha)$ be as in Section 19.9. Observe that

$$n(\alpha, \beta) = n(\beta, \alpha) = \pm 1$$

in this situation, as before. We may as well suppose that $n(\alpha, \beta) = n(\beta, \alpha) = 1$, by replacing $\beta$ with $\sigma_\beta(\beta) = -\beta$ if necessary. Under these conditions, we get that

$$(\sigma_\alpha \circ \sigma_\beta \circ \sigma_\alpha)(\beta) = (\sigma_\alpha \circ \sigma_\beta)(\beta - \alpha) = \sigma_\alpha(-\beta + \alpha + \beta) = \alpha,$$

as desired.

Let $\alpha, \beta \in A$ be given, and let $\sigma$ be an element of $W$ such that

$$(\sigma(\alpha), \beta) \neq 0.$$

If $\sigma(\alpha)$ and $\beta$ are proportional, then

$$\|\alpha\|^2/\|\beta\|^2 = \|\sigma(\alpha)\|^2/\|\beta\|^2 = 1, 4, \text{ or } 1/4.$$
If \( \sigma(\alpha) \) and \( \beta \) are not proportional, then
\[
(20.7.8) \quad \|\alpha\|^2/\|\beta\|^2 = \|\sigma(\alpha)\|^2/\|\beta\|^2 = 1, 2, 3, 1/2, \text{ or } 1/3,
\]
as in Section 19.9. In particular, if \( A \) is reduced, then the only possible ratios are as in (20.7.8). In this case, one can check that there are at most two possible values for the norms of elements of \( A \), as in the first part of Lemma C on p53 of [13]. If \( A \) is not reduced, then one can verify that there are at most three possible values for the norms of elements of \( A \). In order to have three different values for the norms of elements of \( A \), the ratio in (20.7.8) can only be 1, 2, or 1/2 when \( \sigma(\alpha) \) and \( \beta \) are not proportional.

If \( \alpha \in A \), then put \( \tilde{\alpha} = 2\alpha/\|\alpha\|^2 \), as in (19.7.9), so that
\[
(20.7.9) \quad \|\tilde{\alpha}\| = 2/\|\alpha\|.
\]
Remember that \( \hat{A} = \{\tilde{\alpha} : \alpha \in A\} \) is the inverse system of \( A \) in \( V \) with respect to \( (\cdot, \cdot) \), as in Section 19.8. It is easy to see that the irreducibility of \( A \) in \( V \) implies that \( \hat{A} \) is irreducible in \( V \) too. If the elements of \( A \) have the same norm, then \( \hat{A} \) has the same property, by (20.7.9). In this case, \( \hat{A} \) is isomorphic to \( A \), using a dilation on \( V \). Otherwise, the number of distinct values of the norms of elements of \( \hat{A} \) is the same as for \( A \). This corresponds to part of Exercise 11 on p55 of [13].

If \( (\cdot, \cdot) \) is another inner product on \( V \), then there is a unique linear mapping \( T \) from \( V \) into itself such that
\[
(20.7.10) \quad (u, v) = (T(u), v)
\]
for every \( u, v \in V \). More precisely, \( T \) is self-adjoint and positive definite with respect to \( (\cdot, \cdot) \), because of the symmetry and positivity properties of \( (\cdot, \cdot) \) on \( V \). If \( (\cdot, \cdot) \) is also invariant under \( W \), then one can check that \( T \) commutes with the elements of \( W \), because \( (\cdot, \cdot) \) is invariant under \( W \) too. It is well known that there is a basis for eigenvectors of \( T \) in \( V \) that is orthonormal with respect to \( (\cdot, \cdot) \), because \( T \) is self-adjoint with respect to \( (\cdot, \cdot) \). The eigenspaces of \( T \) are mapped into themselves by elements of \( W \), because \( T \) commutes with elements of \( W \). It follows that \( T \) is a constant multiple of the identity mapping on \( V \), because \( W \) acts irreducibly on \( V \). This means that \( (\cdot, \cdot) \) is a constant multiple of \( (\cdot, \cdot) \) on \( V \).

Suppose that \( A \) is reduced, so that there are at most two distinct values of the norms of elements of \( A \). If there are two distinct values for the norms of the elements of \( A \), then the elements of \( A \) whose norm is the larger value are called \textit{long roots}, and the elements of \( A \) whose norm is the smaller value are called \textit{short roots}, as on p53 of [13]. If the elements of \( A \) have the same norm, then it is customary to refer to them as long roots, as in [13].

### 20.8 Maximal roots

Let \( V \) be a vector space over the real numbers of positive finite dimension, let \( A \) be a root system in \( V \), and let \( (\cdot, \cdot) \) be an inner product on \( V \) that is invariant
under the Weyl group of $A$. Also let $B$ be a base for $A$, so that every element of $A$ can be expressed as a linear combination of elements of $B$ with integer coefficients, and where the coefficients are either all greater than or equal to 0, or all less than or equal to 0.

An element $\alpha$ of $A$ is said to be maximal with respect to $B$ if there is no element of $A$ that can be expressed as the sum of $\alpha$ and some elements of $B$, which need not be distinct. Equivalently, this means that if $\gamma \in A$, and each of the coefficients of $\alpha$ associated to elements of $B$ is less than or equal to the corresponding coefficient of $\gamma$, then $\alpha = \gamma$. In particular, $\alpha$ is maximal when the sum of its coefficients associated to elements of $B$ is maximal, as an element of $\mathbb{Z}$.

Suppose from now on in this section that $A$ is irreducible, and let $\alpha_0$ be an element of $A$ that is maximal with respect to $B$. It is easy to see that the coefficients of $\alpha_0$ with respect to elements of $B$ are nonnegative, by maximality. Let $B_0$ be the set of $\beta \in B$ such that the coefficient of $\alpha_0$ with respect to $\beta$ is positive, so that the coefficients of $\alpha_0$ with respect to $\beta \in B \setminus B_0$ are equal to 0. Note that $B_0 \neq \emptyset$, because $\alpha_0 \neq 0$. We would like to show that

\[(20.8.1) \quad B_0 = B,\]

which is part of Lemma A at the bottom of p52 of [13].

If $\beta$ and $\gamma$ are distinct elements of $B$, then $\langle \beta, \gamma \rangle \leq 0$, as in (19.11.8). In particular, this holds when $\beta \in B_0$ and $\gamma \in B \setminus B_0$. If $B_0 \neq B$, then there are $\beta \in B_0$ and $\gamma \in B \setminus B_0$ such that $\langle \beta, \gamma \rangle \neq 0$, because $A$ is irreducible. It follows that $\langle \alpha_0, \gamma \rangle < 0$ under these conditions, by the definition of $B_0$. Observe that $\alpha_0$ is not proportional to $\gamma$ in this situation. This implies that $\alpha_0 + \gamma \in A$, as in Section 19.9. This contradicts the maximality of $\alpha_0$ with respect to $B$, so that (20.8.1) holds.

Next, we would like to check that

\[(20.8.2) \quad \langle \alpha_0, \beta \rangle \geq 0\]

for every $\beta \in B$, which is another part of Lemma A at the bottom of p52 of [13]. If the dimension of $V$ is one, then $B$ has only one element, $\alpha_0$ is a positive integer multiple of that element of $B$, and (20.8.2) follows. Suppose now that the dimension of $V$ is at least two. Let $\beta \in B$ be given, and note that $\alpha_0$ is not proportional to $\beta$, because all of the coefficients of $\alpha_0$ with respect to elements of $B$ are positive. If $\langle \alpha_0, \beta \rangle < 0$, then $\alpha_0 + \beta \in A$, as in Section 19.9 again. This would contradict the maximality of $\alpha_0$ with respect to $B$, so that (20.8.2) holds. We also have that

\[(20.8.3) \quad \langle \alpha_0, \beta \rangle > 0\]

for at least one $\beta \in B$, because $V$ is spanned by $B$.

We would like to show that $\alpha_0$ is unique, as in Lemma A at the bottom of p52 of [13]. Let $\alpha_1$ be another element of $A$ that is maximal with respect to $B$. This implies that the coefficients of $\alpha_1$ with respect to all elements of $B$ are positive, as before. It follows that

\[(20.8.4) \quad \langle \alpha_0, \alpha_1 \rangle > 0,\]
by (20.8.2) and (20.8.3). If \( \alpha_0 \) and \( \alpha_1 \) are proportional, then it is easy to see that \( \alpha_0 = \alpha_1 \), as desired. Otherwise, if \( \alpha_0 \) and \( \alpha_1 \) are not proportional, then \( \alpha_0 - \alpha_1 \in A \), as in Section 19.9. If the coefficients of \( \alpha_0 - \alpha_1 \) with respect to the elements of \( B \) are all nonnegative, then we get a contradiction with the maximality of \( \alpha_1 \). Similarly, if the coefficients of \( \alpha_0 - \alpha_1 \) are all less than or equal to 0, then we get a contradiction with the maximality of \( \alpha_0 \).

If \( \alpha \in A \) and \( \beta \in B \), then

\[
(20.8.5) \quad \text{the coefficient of } \alpha \text{ associated to } \beta \text{ is less than or equal to the coefficient of } \alpha_0 \text{ with respect to } \beta.
\]

To see this, one can first find an element \( \alpha_1 \) of \( A \) that is maximal with respect to \( B \) such that the coefficient of \( \alpha \) associated to \( \beta \) is less than or equal to the coefficient of \( \alpha_1 \) associated to \( \beta \) for every \( \beta \in B \). One can choose \( \alpha_1 \in A \) so that the sum of its coefficients associated to elements of \( B \) is maximal, among elements of \( A \) whose coefficients are greater than or equal to the corresponding coefficients of \( \alpha \). The uniqueness of \( \alpha_0 \) implies that \( \alpha_0 = \alpha_1 \), so that (20.8.5) holds. This is implicit in an argument near the top of p54 of [13]. It follows that the sum of the coefficients of \( \alpha \) associated to elements of \( B \) is less than or equal to the sum of the coefficients of \( \alpha_0 \), with equality only when \( \alpha = \alpha_0 \). This is another part of Lemma A at the bottom of p52 of [13].

### 20.9 More on maximality

Let \( V \) be a vector space over the real numbers of positive finite dimension, and let \( A \) be a root system in \( V \). Also let \( (\cdot, \cdot) \) be an inner product on \( V \) that is invariant under the Weyl group \( W \) of \( A \), and let \( B \) be a base for \( A \). Suppose for the moment that \( A \) is reduced. If \( \tau \) is a linear functional on \( V \), then there is an element \( \sigma \) of \( W \) such that \( \tau(\sigma(\beta)) \geq 0 \) for every \( \beta \in B \), as in Section 19.14.

Let \( \alpha \in V \) be given, so that

\[
(20.9.1) \quad \tau_\alpha(v) = (v, \alpha)
\]

defines a linear functional on \( V \). As before, there is a \( \sigma \in W \) such that

\[
(20.9.2) \quad (\sigma(\beta), \alpha) = \tau_\alpha(\sigma(\beta)) \geq 0
\]

for every \( \beta \in B \). Equivalently, this means that

\[
(20.9.3) \quad (\beta, \sigma^{-1}(\alpha)) \geq 0
\]

for every \( \beta \in B \), because \( \sigma \) is an orthogonal transformation on \( V \) with respect to \( (\cdot, \cdot) \). This corresponds to the formulation in part (a) of the theorem on p51 and Exercise 14 on p55 of [13].

Suppose now that \( A \) is irreducible as well as reduced, and let \( \alpha_0 \in A \) be maximal with respect to \( B \), as in the previous section. We would like to show that

\[
(20.9.4) \quad ||\alpha_0|| \geq ||\alpha||
\]
20.9. MORE ON MAXIMALITY

for every $\alpha \in A$, where $\| \cdot \|$ is the norm on $V$ associated to $(\cdot, \cdot)$, as in Lemma D on p53 of [13]. We may as well suppose that

$$(20.9.5) \quad (\beta, \alpha) \geq 0$$

for every $\beta \in B$, by replacing $\alpha$ with $\sigma^{-1}(\alpha)$ for $\sigma \in W$ as in (20.9.3). Note that the coefficients of $\alpha$ associated to elements of $B$ are less than or equal to the corresponding coefficients of $\alpha_0$, as in (20.8.5). Combining this with (20.9.5), we get that

$$(20.9.6) \quad (\alpha_0 - \alpha, \alpha) \geq 0,$$

which is the same as saying that

$$(20.9.7) \quad (\alpha_0, \alpha) \geq (\alpha, \alpha).$$

This implies that

$$(20.9.8) \quad \| \alpha \|^2 = (\alpha, \alpha) \leq \| \alpha_0 \| \| \alpha \|,$$

by the Cauchy–Schwarz inequality. It is easy to obtain (20.9.4) from this, as desired.

Let $A$ be a root system in $V$ again, that is not necessarily reduced. If $\alpha \in A$, then put $\hat{\alpha} = 2\alpha/\|\alpha\|^2$, as in (19.7.9), and remember that $\hat{A} = \{\hat{\alpha} : \alpha \in A\}$ is the inverse system of $A$ in $V$ with respect to $(\cdot, \cdot)$, as in Section 19.8. Let $B$ be a base for $A$, as before, and let $B_1, B_2$ be the sets of $\beta \in B$ such that $2\beta \not\in A$ and $2\beta \in A$, respectively, as in Section 19.13. Remember that

$$(20.9.9) \quad \hat{B} = \{\hat{\beta} : \beta \in B_1\} \cup \{\hat{\beta}/2 : \beta \in B_2\}$$

is a base for $\hat{A}$. If $\alpha, \gamma \in A$ satisfy

$$(20.9.10) \quad \|\alpha\| = \|\gamma\|,$$

then one can check that the coefficients of $\alpha$ associated to elements of $B$ are less than or equal to the corresponding coefficients of $\gamma$ if and only if the coefficients of $\hat{\alpha}$ associated to elements of $\hat{B}$ are less than or equal to the corresponding coefficients of $\hat{\gamma}$.

Suppose that $A$ is reduced, so that $B_2 = \emptyset$, and $\hat{B}$ is the same as

$$(20.9.11) \quad \hat{B} = \{\hat{\beta} : \beta \in B\}.$$ Suppose that $A$ is irreducible as well, which implies that $\hat{A}$ is irreducible too, as in Section 20.7. Remember that there are at most two distinct values for the norms of the elements of $A$, as in Section 20.7. Let us suppose that there are two distinct values for the norms of elements of $A$, which means that there are two distinct values for the norms of elements of $A$.

Let $\hat{\gamma}_0$ be a maximal root in $\hat{A}$, as in the previous section. This implies that $\hat{\gamma}_0$ is a long root in $\hat{A}$, as before. It follows that the corresponding element $\gamma_0$ of $A$ is a short root in $A$. If $\alpha \in A$, then the coefficients of $\hat{\alpha}$ associated to elements of $\hat{B}$ are less than or equal to the corresponding coefficients of $\hat{\gamma}_0$, as in (20.8.5). If $\alpha$ is a short root in $A$, then $\|\alpha\| = \|\gamma_0\|$, and it follows that the coefficients of $\alpha$ associated to elements of $B$ are less than or equal to the coefficients of $\gamma_0$, as before. Note that $\gamma_0$ is uniquely determined by this property. This corresponds to part of Exercise 11 on p55 of [13].
20.10 Root systems of type $A_n$

Let $n$ be a positive integer, and let $\mathbb{R}^{n+1}$ be the usual space of $(n + 1)$-tuples of real numbers. The standard inner product is defined on $\mathbb{R}^{n+1}$ by

\begin{equation}
(u, v) = \sum_{j=1}^{n+1} u_j v_j,
\end{equation}

as usual. Let $e_1, \ldots, e_{n+1}$ be the standard basis vectors in $\mathbb{R}^{n+1}$, so that the $j$th coordinate of $e_l$ is equal to 1 when $j = l$, and to 0 otherwise. Of course, the set $\mathbb{Z}^{n+1}$ of $(n + 1)$-tuples of integers is a subgroup of $\mathbb{R}^{n+1}$, as a commutative group with respect to addition. Put

\begin{equation}
V = \left\{ v \in \mathbb{R}^{n+1} : \sum_{j=1}^{n+1} v_j = 0 \right\},
\end{equation}

which is the hyperplane in $\mathbb{R}^{n+1}$ orthogonal to $e_1 + \cdots + e_{n+1}$.

Consider

\begin{equation}
A = \{ \alpha \in V \cap \mathbb{Z}^{n+1} : (\alpha, \alpha) = 2 \}.
\end{equation}

Equivalently, it is easy to see that $A$ consists exactly of the vectors of the form $e_j - e_l$, where $1 \leq j, l \leq n + 1$ and $j \neq l$. Note that $A$ is a finite set of nonzero elements of $V$ whose linear span is equal to $V$.

Suppose for the moment that $2 \in \mathbb{R}^{n+1}$ satisfies $(\alpha, \alpha) = 2$. In this case, the reflection on $\mathbb{R}^{n+1}$ associated to $\alpha$ with respect to the standard inner product is given by

\begin{equation}
\sigma_\alpha(v) = v - (v, \alpha) \alpha
\end{equation}

for every $v \in \mathbb{R}^{n+1}$. If $\alpha \in V$, then $\sigma_\alpha$ maps $V$ onto itself, as in Section 20.6. If $\alpha \in \mathbb{Z}^{n+1}$, then $\sigma_\alpha(\mathbb{Z}^{n+1}) \subseteq \mathbb{Z}^{n+1}$, and hence

\begin{equation}
\sigma_\alpha(\mathbb{Z}^{n+1}) = \mathbb{Z}^{n+1},
\end{equation}

because $\sigma_\alpha$ is its own inverse. It follows that $\sigma_\alpha(A) = A$ when $\alpha \in A$, because $\sigma_\alpha$ is an orthogonal transformation on $\mathbb{R}^{n+1}$. Of course, $(v, \alpha) \in \mathbb{Z}$ when $\alpha, v \in \mathbb{Z}^{n+1}$, so that $A$ defines a root system in $V$. This root system is said to be of type $A_n$. Clearly $A$ is reduced, as a root system in $V$.

Suppose that $\alpha = e_j - e_l$ for some $j, l \in \{1, \ldots, n+1\}$, with $j \neq l$. If $v \in \mathbb{R}^{n+1}$, then

\begin{equation}
\sigma_\alpha(v) = v - (v, e_l) (e_j - e_l),
\end{equation}

which interchanges the $j$th and $l$th coordinates of $v$, without affecting the other coordinates of $v$. It follows that the Weyl group of $A$ consists of the linear mappings on $V$ obtained from permuting the coordinates of elements of $V$.

Remember that the inverse system $\tilde{A}$ of $A$ in $V$ with respect to $(\cdot, \cdot)$ is defined in (19.8.4). In this case, $\tilde{A} = A$. 


20.10. ROOT SYSTEMS OF TYPE $A_N$

If $\alpha, \beta \in A$ and $n(\alpha, \beta) \in \mathbb{Z}$ is as in Sections 19.9 and 20.2, then

\begin{equation}
(20.10.7) \quad n(\alpha, \beta) = (\alpha, \beta)
\end{equation}

in this situation, as in (20.2.2). Thus

\begin{equation}
(20.10.8) \quad n(\alpha, \beta) n(\beta, \alpha) = (\alpha, \beta)^2.
\end{equation}

If $\alpha = \pm \beta$, then (20.10.7) is $\pm 2$, and (20.10.8) is 4. Suppose that $\alpha = e_{j_1} - e_{j_2}$ and $\beta = e_{l_1} - e_{l_2}$ for some $j_1, j_2, l_1, l_2 \in \{1, \ldots, n+1\}$ with $j_1 \neq j_2$, $l_1 \neq l_2$. If $\{j_1, j_2\} \cap \{l_1, l_2\}$ has exactly one element, then (20.10.7) is $\pm 1$, and (20.10.8) is 1. If $\{j_1, j_2\} \cap \{l_1, l_2\} = \emptyset$, then (20.10.7) and (20.10.8) are equal to 0.

Put

\begin{equation}
(20.10.9) \quad B = \{e_j - e_{j+1} : j = 1, \ldots, n\},
\end{equation}

which is a subset of $A$. It is easy to see that $B$ is a basis for $V$, as a vector space over the real numbers. More precisely, one can check that $B$ is a base for $A$ as a root system in $V$. Indeed, if $1 \leq j < l \leq n+1$, then

\begin{equation}
(20.10.10) \quad e_j - e_l = \sum_{r=j}^{l-1} (e_r - e_{r+1}).
\end{equation}

In particular, $e_1 - e_{n+1}$ is the unique maximal element of $A$ with respect to $B$, as in Section 20.8.

Let $j, l \in \{1, \ldots, n\}$ be given, so that $e_j - e_{j+1}$ and $e_l - e_{l+1}$ are elements of (20.10.9). The corresponding entry of the Cartan matrix of $A$ with respect to $B$ is given by

\begin{equation}
(20.10.11) \quad n(e_j - e_{j+1}, e_l - e_{l+1}) = (e_j - e_{j+1}, e_l - e_{l+1}).
\end{equation}

This is symmetric in $j$ and $l$, and equal to 2 when $j = l$. As before, (20.10.11) is equal to $-1$ when $|j - l| = 1$, and to 0 when $|j - l| \geq 2$.

Similarly, let us consider the Coxeter graph of $A$ with respect to $B$, as in Section 20.3. If $j, l \in \{1, \ldots, n\}$ and $j \neq l$, then the number of edges between $e_j - e_{j+1}$ and $e_l - e_{l+1}$ in the Coxeter graph is the square of (20.10.11). This is equal to 1 when $|j - l| = 1$, and to 0 when $|j - l| \geq 2$. Note that the Coxeter graph is connected, so that $A$ is irreducible.

Let $T$ be a one-to-one linear mapping from $V$ onto itself such that $T(A) = A$. This implies that

\begin{equation}
(20.10.12) \quad (T(\alpha), T(\beta)) = (\alpha, \beta)
\end{equation}

for every $\alpha, \beta \in A$, by (20.2.13) and (20.10.7). It follows that (20.10.12) holds for every $\alpha, \beta \in V$, because $V$ is spanned by $A$.

Suppose that $T$ also satisfies $T(B) = B$. If $n = 1$, then $T$ is the identity mapping on $V$. If $n \geq 2$, then there is another such mapping $T$ on $V$, which sends $e_j - e_{j+1}$ to $e_{n-j} - e_{n-j+1}$ for each $j = 1, \ldots, n$. Equivalently, $T$ is the restriction to $V$ of the mapping on $\mathbb{R}^{n+1}$ that sends $v = (v_1, \ldots, v_{n+1})$ to $-(v_{n+1}, \ldots, v_1)$.

This corresponds to some of the remarks on p64 and in Tables 1 and 2 on p66 of [13], and on p39-40 of [23].
20.11 Types $B_n$ and $C_n$

Let $n$ be a positive integer, and let us take $V = \mathbb{R}^n$, with its standard inner product $(\cdot, \cdot)$ and basis $e_1, \ldots, e_n$. Consider

\[(20.11.1) \hspace{1cm} A_1 = \{\alpha \in \mathbb{Z}^n : (\alpha, \alpha) = 1 \text{ or } 2\}.\]

Equivalently, $A_1$ consists of vectors of the form $\pm e_j$, $1 \leq j \leq n$, and $\pm e_j \pm e_l$, $1 \leq j \neq l \leq n$. Similarly, put

\[(20.11.2) \hspace{1cm} A_2 = \{\pm 2e_j : 1 \leq j \leq n\} \cup \{\pm e_j \pm e_l : 1 \leq j \neq l \leq n\},\]

which is a subset of $\mathbb{Z}^n$ as well. Clearly $A_1$ and $A_2$ are finite sets of nonzero elements of $V$ that span $V$. Note that the squares of the norms of the elements of $A_2$ are either 2 or 4.

If $\alpha$ is a nonzero element of $V$, then let $\sigma_\alpha$ be the reflection on $V$ associated to $\alpha$ with respect to $(\cdot, \cdot)$, as usual. Suppose for the moment that $\alpha$ is a nonzero multiple of $e_j$ for some $j = 1, \ldots, n$. In this case, $\sigma_\alpha = e_j$ multiplies the $j$th coordinate of $v \in V$ by $-1$, and leaves the other coordinates of $v$ unchanged. In particular, $\sigma_\alpha$ maps $\mathbb{Z}^n$, $A_1$, and $A_2$ onto themselves. Equivalently,

\[(20.11.3) \hspace{1cm} \sigma_\alpha(v) - v = -2v_j e_j\]

for every $v \in V$. If $\alpha = \pm e_j$ or $\pm 2e_j$ and $v \in \mathbb{Z}^n$, then it follows that $\sigma_\alpha(v) - v$ is an integer multiple of $\alpha$. Thus this holds for every $v \in A_1$ or $A_2$.

If $\alpha \in V$ satisfies $(\alpha, \alpha) = 2$, then $\sigma_\alpha(v) = v - (v, \alpha)\alpha$ for every $v \in V$, as before. If we also have that $\alpha \in \mathbb{Z}^n$, then $\sigma_\alpha$ maps $\mathbb{Z}^n$ into itself, and hence onto itself. In this case, if $v \in \mathbb{Z}^n$, then $(v, \alpha) \in \mathbb{Z}$, so that $\sigma_\alpha(v) - v$ is an integer multiple of $\alpha$. If $\alpha = e_j - e_l$, $1 \leq j \neq l \leq n$, then $\sigma_\alpha(v)$ interchanges the $j$th and $l$th coordinates of $v$, without affecting the other coordinates, as in the previous section. This implies that $\sigma_\alpha$ maps $A_1$ and $A_2$ onto themselves.

Suppose that $\alpha = \pm(e_j + e_l)$ for some $1 \leq j \neq l \leq n$, so that $(\alpha, \alpha) = 2$. In this case,

\[(20.11.4) \hspace{1cm} \sigma_\alpha(v) = v - (v_j + v_l)(e_j + e_l)\]

for every $v \in V$. This interchanges the $j$th and $l$th coordinates and multiplies them by $-1$, while leaving the other coordinates unchanged. Thus $\sigma_\alpha$ maps $A_1$ and $A_2$ onto themselves, so that $A_1$ and $A_2$ are root systems in $V$. More precisely, $A_1$ is said to be of type $B_n$, and $A_2$ is said to be of type $C_n$.

Consider

\[(20.11.5) \hspace{1cm} A_3 = A_1 \cup A_2 = \{\pm e_j : 1 \leq j \leq n\} \cup \{\pm 2e_j : 1 \leq j \leq n\} \cup \{\pm e_j \pm e_l : 1 \leq j \neq l \leq n\}.\]

If $\alpha \in A_3$, then $\sigma_\alpha$ maps $A_1$ and $A_2$ onto themselves, as in the previous paragraphs, and hence $\sigma_\alpha$ maps $A_3$ onto itself. We also have that $\sigma_\alpha(v) - v$ is an integer multiple of $\alpha$ for every $v \in \mathbb{Z}^n$, as before, and in particular for $v \in A_3$. 

This means that $A_3$ is a root system in $V$ as well, which is said to be of type $BC_n$. Note that $A_1$ and $A_2$ are reduced, and that $A_3$ is not reduced.

If $\alpha$ is a nonzero element of $V$, then put $\widehat{\alpha} = 2\alpha/(\alpha, \alpha)$, as in (19.7.9). Observe that

\begin{equation}
(20.11.6) \quad A_2 = \widehat{A}_1 = \{\widehat{\alpha} : \alpha \in A_1\},
\end{equation}

which is the inverse system of $A_1$ in $V$ with respect to $(\cdot, \cdot)$, as in Section 19.8. Of course, this means that $A_1 = \widehat{A}_2$ is the inverse system of $A_2$ in $V$ with respect to $(\cdot, \cdot)$. It follows that $\widehat{A}_3 = A_3$.

The Weyl groups of $A_1$, $A_2$, and $A_3$ are the same, because they are generated by the same reflections on $V$. This group consists of linear mappings on $V$ that permute the coordinates of elements of $V$, while multiplying any collection of the coordinates by $-1$.

Put

\begin{equation}
(20.11.7) \quad B_1 = \{e_j - e_{j+1} : 1 \leq j \leq n - 1\} \cup \{e_n\},
\end{equation}

which is interpreted as being $\{e_1\}$ when $n = 1$. It is easy to see that this is a basis for $V$ as a vector space over $\mathbb{R}$, and one can check that it is a base for $A_1$, as a root system in $V$. More precisely, if $1 \leq j < n$, then

\begin{equation}
(20.11.8) \quad e_j = \sum_{r=j}^{n-1} (e_r - e_{r+1}) + e_n
\end{equation}

and

\begin{equation}
(20.11.9) \quad e_j + e_n = \sum_{r=j}^{n-1} (e_r - e_{r+1}) + 2e_n.
\end{equation}

If $1 \leq j < l \leq n$, then

\begin{equation}
(20.11.10) \quad e_j - e_l = \sum_{r=j}^{l-1} (e_r - e_{r+1}).
\end{equation}

If $1 \leq j < l < n$, then

\begin{equation}
(20.11.11) \quad e_j + e_l = \sum_{r=j}^{l-1} (e_r - e_{r+1}) + \sum_{r=l}^{n-1} 2(e_r - e_{r+1}) + 2e_n.
\end{equation}

Similarly, put

\begin{equation}
(20.11.12) \quad B_2 = \{e_j - e_{j+1} : 1 \leq j \leq n - 1\} \cup \{2e_n\},
\end{equation}

which is interpreted as being $\{2e_1\}$ when $n = 1$. This is a basis for $V$ as a vector space over $\mathbb{R}$, and one can verify that it is a base for $A_2$ as a root system in $V$. Indeed, if $1 \leq j < n$, then

\begin{equation}
(20.11.13) \quad 2e_j = \sum_{r=j}^{n-1} 2(e_r - e_{r+1}) + 2e_n.
\end{equation}
and $e_j + e_n$ can be expressed as in (20.11.9). If $1 \leq j < l \leq n$, then $e_j - e_l$ can be expressed as in (20.11.10), and $e_j + e_l$ can be expressed as in (20.11.11) when $l < n$. Note that $B_1$ is a base for $A_3$ as well, using (20.11.13).

If $n = 1$, then $A_1 = \{ \pm e_1 \}$ and $A_2 = \{ \pm 2e_1 \}$ are isomorphic to each other, and to the root system discussed in the previous section, with $n = 1$. If $n = 2$, then one can use the bases $B_1$ and $B_2$ to show that $A_1$ is isomorphic to $A_2$.

Let $\alpha$ and $\beta$ be elements of $A_1$, $A_2$, or $A_3$, so that $n(\alpha, \beta) \in \mathbb{Z}$ can be defined as in Section 19.9. In fact,

$$n(\alpha, \beta) = 2 (\alpha, \beta) (\beta, \beta)^{-1}, \tag{20.11.14}$$

as before, so that

$$n(\alpha, \beta) n(\beta, \alpha) = 2 (\alpha, \beta)^2 (\alpha, \alpha)^{-1} (\beta, \beta)^{-1}. \tag{20.11.15}$$

If $1 \leq j, l \leq n - 1$, then

$$n(e_j - e_{j+1}, e_l - e_{l+1}) = (e_j - e_{j+1}, e_l - e_{l+1}) \tag{20.11.16}$$

is symmetric in $j$ and $l$, and equal to 2 when $j = l$. This is equal to $-1$ when $|j - l| = 1$, and to 0 when $|j - l| \geq 2$, as before. It follows that

$$n(e_j - e_{j+1}, e_l - e_{l+1}) n(e_l - e_{l+1}, e_j - e_{j+1}) \tag{20.11.17}$$

is equal to 1 when $|j - l| = 1$, and to 0 when $|j - l| \geq 2$.

If $1 \leq j \leq n - 1$, then

$$n(e_j - e_{j+1}, e_n) = 2 (e_j - e_{j+1}, e_n), \tag{20.11.18}$$

which is 0 when $j < n - 1$, and $-2$ when $j = n - 1$. Similarly,

$$n(e_n, e_j - e_{j+1}) = (e_n, e_j - e_{j+1}), \tag{20.11.19}$$

which is 0 when $j < n - 1$ and $-1$ when $j = n - 1$. Thus

$$n(e_j - e_{j+1}, e_n) n(e_n, e_j - e_{j+1}) \tag{20.11.20}$$

is equal to 0 when $j < n - 1$, and to 2 when $j = n - 1$. Using this and the remarks in the preceding paragraph, one can determine the Cartan matrix and Coxeter graph of $A_1$ with respect to $B_1$. These are the same as for $A_3$ with respect to $B_1$.

If $1 \leq j \leq n - 1$ again, then

$$n(e_j - e_{j+1}, 2e_n) = (e_j - e_{j+1}, e_n), \tag{20.11.21}$$

which is 0 when $j < n - 1$ and $-1$ when $j = n - 1$. We also have that

$$n(2e_n, e_j - e_{j+1}) = 2 (e_n, e_j - e_{j+1}), \tag{20.11.22}$$
which is 0 when \( j < n - 1 \), and \(-2\) when \( j = n - 1 \). This implies that

\[
(20.11.23) \quad n(e_j - e_{j+1}, 2e_n) n(2e_n, e_j - e_{j+1})
\]

equal to 0 when \( j < n - 1 \), and to 2 when \( j = n - 1 \). This can be used to determine the Cartan matrix and Coxeter graph of \( A_2 \) with respect to \( B_2 \), in combination with the earlier remarks about (20.11.16) and (20.11.17). Note that the Coxeter graphs of \( A_1 \), \( A_2 \), and \( A_3 \) are connected, so that these root systems are irreducible.

If \( T \) is an automorphism of \( A_1 \) or \( A_3 \) and \( T(B_1) = B_1 \), then one can check that \( T \) is the identity mapping on \( V \), using the Cartan matrix. Similarly, if \( T \) is an automorphism of \( A_2 \) and \( T(B_2) = B_2 \), then \( T \) is the identity mapping on \( V \).

If \( n \geq 2 \), then the maximal root in \( A_1 \) with respect to \( B_1 \) is \( e_1 + e_2 \), which can be expressed in terms of \( B_1 \) as in (20.11.9) when \( n = 2 \), and (20.11.11) when \( n > 2 \). The maximal short root in \( A_1 \) with respect to \( B_1 \) is \( e_1 \), which can be expressed in terms of \( B_1 \) as in (20.11.8). The maximal root in \( A_2 \) with respect to \( B_2 \) is \( 2e_1 \), which can be expressed in terms of \( B_2 \) as in (20.11.13). The maximal short root in \( A_2 \) with respect to \( B_2 \) is \( e_1 + e_2 \) when \( n \geq 2 \), which can be expressed in terms of \( B_2 \) as in (20.11.9) when \( n = 2 \), and (20.11.11) when \( n > 2 \).

This corresponds to some of the remarks on p64 and Tables 1 and 2 on p66 of [13], and on p39-40 of [23], as before. Root systems of type \( BC_n \) are mentioned in Exercise 3 on p66 of [13], and on p41 of [23].

### 20.12 Type \( D_n \)

Let \( n \geq 2 \) be an integer, and let us take \( V = \mathbb{R}^n \) again, with its standard inner product \( (\cdot, \cdot) \) and basis \( e_1, \ldots, e_n \). Consider

\[
(20.12.1) \quad A = \{ \alpha \in \mathbb{Z}^n : (\alpha, \alpha) = 2 \},
\]

which consists of the vectors of the form \( \pm e_j \pm e_l, 1 \leq j \neq l \leq n \). Of course, \( A \) is a finite set of nonzero elements of \( V \) that spans \( V \).

If \( \alpha \in V \) and \((\alpha, \alpha) = 2\), then the reflection \( \sigma_\alpha \) on \( V \) associated to \( \alpha \) with respect to \((\cdot, \cdot)\) is given by \( \sigma_\alpha(v) = v - (v, \alpha) \alpha \), as before. If \( \alpha \in \mathbb{Z}^n \) too, then \( \sigma_\alpha \) maps \( \mathbb{Z}^n \) onto itself, and hence \( \sigma_\alpha \) maps \( A \) onto itself, because \( \sigma_\alpha \) is an orthogonal transformation on \( V \). We also have that \( \sigma_\alpha(v) - v = (v, \alpha) \alpha \) is an integer multiple of \( \alpha \) when \( v \in \mathbb{Z}^n \) in this situation. Thus \( A \) is a root system in \( V \), which is said to be of type \( D_n \). Observe that \( A \) is reduced, as a root system in \( V \).

If \( \alpha = e_j - e_l \) for some \( 1 \leq j \neq l \leq n \), then \( \sigma_\alpha(v) \) interchanges the \( j \)th and \( l \)th coordinates of \( v \in V \), leaving the other coordinates unchanged, as before. If \( \alpha = \pm(e_j + e_l) \) for some \( 1 \leq j \neq l \leq n \), then \( \sigma_\alpha(v) \) interchanges the \( j \)th and \( l \)th coordinates and multiplies them by \(-1\), leaving the other coordinates unchanged, as in the previous section. The Weyl group of \( A \) consists of linear mappings on \( V \) that permute the coordinates of \( v \in V \), and multiply an even number of the coordinates by \(-1\).
Let $\hat{A}$ be the inverse system of $A$ in $V$ with respect to $(\cdot, \cdot)$, as in Section 19.8. Observe that $\hat{A} = A$ in this situation.

It is easy to see that

$$B = \{e_j - e_{j+1} : 1 \leq j \leq n - 1\} \cup \{e_{n-1} + e_n\}$$

is a basis for $V$, as a vector space over $\mathbb{R}$. One can check that $B$ is a base for $A$, as a root system in $V$. More precisely, if $1 \leq j < l \leq n$, then

$$e_j - e_l = \sum_{r=j}^{l-1} (e_r - e_{r+1}),$$

as usual. If $1 \leq j < n - 1$, then

$$e_j + e_n = \sum_{r=j}^{n-2} (e_r - e_{r+1}) + (e_{n-1} + e_n).$$

If $1 \leq j < l \leq n - 1$, then

$$e_j + e_l = \sum_{r=j}^{l-1} (e_r - e_{r+1}) + \sum_{r=l}^{n-2} 2(e_r - e_{r+1}) + (e_{n-1} - e_n) + (e_{n-1} + e_n),$$

where the second sum on the right is interpreted as being equal to 0 when $l = n - 1$.

If $\alpha, \beta \in A$, then $n(\alpha, \beta) \in \mathbb{Z}$ can be defined as in Section 19.9, and is equal to $(\alpha, \beta)$ in this situation. If $1 \leq j, l \leq n - 1$, then

$$n(e_j - e_{j+1}, e_l - e_{l+1}) = (e_j - e_{j+1}, e_l - e_{l+1})$$

is equal to 2 when $j = l$, to $-1$ when $|j - l| = 1$, and to 0 when $|j - l| \geq 2$, as before. Thus

$$n(e_j - e_{j+1}, e_l - e_{l+1}) n(e_l - e_{l+1}, e_j - e_{j+1})$$

is equal to 1 when $|j - l| = 1$, and to 0 when $|j - l| \geq 2$. If $1 \leq j \leq n - 1$, then

$$n(e_j - e_{j+1}, e_{n-1} + e_n) = (e_j - e_{j+1}, e_{n-1} + e_n)$$

is equal to 0 when $j \leq n - 3$ and when $j = n - 1$, and equal to $-1$ when $j = n - 2$. It follows that

$$n(e_j - e_{j+1}, e_{n-1} + e_n) n(e_{n-1} + e_n, e_j - e_{j+1})$$

is equal to 0 when $j \leq n - 3$ and $j = n - 1$, and to 1 when $j = n - 1$. This can be used to determine Cartan matrix and Coxeter graph of $A$ with respect to $B$. If $n \geq 3$, then the Coxeter graph is connected, so that $A$ is irreducible in
20.13. COMPLEX NUMBERS

V. In this case, the maximal root in A with respect to B is \( e_1 + e_2 \), which can expressed in terms of B as in (20.12.5).

If \( n = 2 \), then A is isomorphic to the sum of two root systems of rank one. If \( n = 3 \), then A is isomorphic to the corresponding root system in Section 20.10. Thus one often restricts one’s attention to \( n \geq 4 \).

As usual, one can look for automorphisms \( T \) of A that map B onto itself, other than the identity mapping. One way to do this is to interchange \( e_{n-1} + e_n \) with \( e_{n-1} \) while fixing the other elements of B. If \( n = 4 \), then one can permute any of the basis elements \( e_1 - e_2, e_3 - e_4, \) and \( e_3 + e_4 \).

This corresponds to some remarks on p64 and Tables 1 and 2 on p66 of [13] again, and on p39-40 of [23].

20.13 Complex numbers

Let V be a vector space over the real numbers of positive finite dimension, and let \((\cdot, \cdot)\) be an inner product on V. If \( \alpha \) is a nonzero element of V, then let \( \sigma_\alpha \) be the reflection on V associated to \( \alpha \) with respect to \((\cdot, \cdot)\). Let \( \beta \) be another nonzero element of V. If \( \alpha \) and \( \beta \) are proportional in V, then \( \sigma_\alpha = \sigma_\beta \) on V. Otherwise, if \( \alpha \) and \( \beta \) are not proportional in V, then they span a two-dimensional linear subspace \( V_0 \) of V. Note that \( \sigma_\alpha \) and \( \sigma_\beta \) map \( V_0 \) onto itself, as in Section 20.6. Of course, \( \sigma_\alpha \) and \( \sigma_\beta \) fix every element of the orthogonal complement of \( V_0 \) in V.

The complex plane \( \mathbb{C} \) may be considered as a two-dimensional vector space over the real numbers. More precisely, \( \mathbb{C} \) can be identified with \( \mathbb{R}^2 \) in the usual way, using the real and imaginary parts of a complex number. If \( u, v \in \mathbb{C} \), then put

\[
(20.13.1) \quad (u, v) = \Re(u \overline{v}),
\]

which is the real part of \( u \) times the complex conjugate of \( v \). This is the standard inner product on \( \mathbb{C} \), as a vector space over \( \mathbb{R} \), and which corresponds to the standard inner product on \( \mathbb{R}^2 \). The norm on \( \mathbb{C} \) associated to (20.13.1) is the same as the standard absolute value function \( |\cdot| \) on \( \mathbb{C} \). Note that

\[
(20.13.2) \quad (au, av) = |a|^2 (u, v)
\]

for every \( a, u, v \in \mathbb{C} \). As before, a real-linear mapping from \( \mathbb{C} \) into itself is a linear mapping from \( \mathbb{C} \) into itself, as a vector space over \( \mathbb{R} \). Complex conjugation on \( \mathbb{C} \) is a real-linear mapping from \( \mathbb{C} \) into itself that preserves (20.13.2). This is the same as the reflection on \( \mathbb{C} \), as a vector space over \( \mathbb{R} \), associated to a nonzero imaginary number with respect to (20.13.1).

If \( a \) is a nonzero complex number, then put

\[
(20.13.3) \quad \rho_\alpha(z) = \overline{a(z/a)} = (a/\overline{a}) \overline{z}
\]

for every \( z \in \mathbb{C} \). This defines a real-linear mapping \( \rho_\alpha \) from \( \mathbb{C} \) into itself, with

\[
(20.13.4) \quad \rho_{au} = \rho_a
\]
for every \( t \in \mathbb{R} \setminus \{0\} \). Observe that \( \rho_a \) is an orthogonal transformation on \( \mathbb{C} \), as a vector space over \( \mathbb{R} \), with respect to (20.13.1). If \( \alpha \in \mathbb{C} \setminus \{0\} \), then let \( \sigma_\alpha \) be the reflection on \( \mathbb{C} \), as a vector space over \( \mathbb{R} \), associated to \( \alpha \) with respect to (20.13.1). Observe that

\[
\rho_a = \sigma_{\alpha i}
\]

for every \( \alpha \in \mathbb{C} \setminus \{0\} \). Equivalently,

\[
\sigma_\alpha = \rho_{\alpha i} = -\rho_\alpha
\]

for every \( \alpha \in \mathbb{C} \setminus \{0\} \). Thus

\[
\sigma_\alpha(z) = -\rho_\alpha(z) = -\alpha \overline{(z/\alpha)} = -(\alpha/\pi) \overline{z}
\]

for every \( \alpha \in \mathbb{C} \setminus \{0\} \) and \( z \in \mathbb{C} \).

If \( \alpha, \beta \in \mathbb{C} \setminus \{0\} \) and \( z \in \mathbb{C} \), then

\[
\sigma_\alpha(\sigma_\beta(z)) = -(\alpha/\pi) \overline{\sigma_\beta(z)} = (\alpha/\pi) \overline{(\beta/\pi)} z.
\]

This is related to Exercise 3 on p46 of [13]. This is also related to Weyl groups of rank two root systems, as in Exercise 4 on p46 of [13], and some remarks on p27 of [23].

Let \( \alpha \in \mathbb{C} \setminus \{0\} \) and \( z \in \mathbb{C} \) be given. Observe that

\[
\sigma_\alpha(z) - z = -\alpha \overline{(z/\alpha)} - z = -\alpha \overline{(z/\alpha)} \overline{\alpha} = -2 \text{ Re}(z/\alpha) \alpha = -2 (z, \alpha) |\alpha|^{-2} \alpha.
\]

Of course, the right side corresponds to the usual expression for reflections with respect to an inner product.

### 20.14 Type \( A_2 \)

Some very nice pictures of a root system in the plane that is isomorphic to the one in Section 20.10 with \( n = 2 \) can be found on p44 of [13] and p26 of [23]. This root system can also be described as the set \( A \) of cube roots of unity and their negatives in the complex plane. Remember that points on the unit circle in \( \mathbb{C} \) can be expressed as \( \exp(it) \) with \( t \in \mathbb{R} \), using the complex exponential function. The elements of \( A \) can be expressed as \( \exp(it) \), where \( t \) is an integer multiple of \( \pi/3 \). In particular, \( A \) is a subgroup of the unit circle in \( \mathbb{C} \), as a commutative group with respect to multiplication.

It is well known that

\[
\exp(\pi i/3) = 1/2 + (\sqrt{3}/2) i
\]

and

\[
\exp(2 \pi i/3) = -1/2 + (\sqrt{3}/2) i.
\]
Indeed,

\[(20.14.3) \quad (1/2 + (\sqrt{3}/2) i)^2 = 1/4 - 3/4 + 2 (1/2) (\sqrt{3}/2) i = -1/2 + (\sqrt{3}/2) i \]

and

\[(20.14.4) \quad (1/2 + (\sqrt{3}/2) i)^3 = (1/2 + (\sqrt{3}/2) i)^2 (1/2 + (\sqrt{3}/2) i) = (-1/2 + (\sqrt{3}/2) i) (1/2 + (\sqrt{3}/2) i) = -1/4 - 3/4 = -1. \]

Put \( \alpha = 1 \) and \( \beta = \exp(2 \pi i/3) \), so that \( \alpha + \beta = \exp(\pi i/3) \), as indicated in the pictures on p44 of [13] and p26 of [23]. Thus \( A \) consists of \( \pm \alpha, \pm \beta \), and \( \pm (\alpha + \beta) \). Of course, \( A \) is a finite set of nonzero elements of \( C \), whose linear span in \( C \), as a vector space over \( R \), is equal to \( C \).

If \( \gamma \in C \setminus \{0\} \), then let \( \sigma_\gamma \) be the reflection on \( C \), as a vector space over \( R \), associated to \( \gamma \) with respect to the standard inner product, as before. Thus one should check that \( \sigma_\gamma(A) = A \) for every \( \gamma \in A \). This is clear from the pictures in [13, 23], and can be verified using (20.13.7) as well. One should also check that \( \sigma_\gamma(z) - z \) is an integer multiple of \( \gamma \in A \) for every \( z \in A \), which means that \( 2(z, \gamma) \in Z \) for every \( z \in A \). This can be seen using the expressions for the elements of \( A \) mentioned earlier, or more geometrically.

Clearly \( A \) is reduced, as a root system in the plane. It is easy to see that \( B = \{\alpha, \beta\} \) is a base for \( A \), as in [13, 23]. If we take \( n = 2 \) in Section 20.10, then we have seen that \( \{e_1 - e_2, e_2 - e_3\} \) is a base for the root system discussed there. One can check that these root systems are isomorphic, using a linear mapping which sends one base to the other. Note that the inner product in Section 20.10 corresponds to twice the standard inner product on \( C \) in this way.

### 20.15 Dynkin diagrams

Let \( V \) be a vector space over the real numbers of positive finite dimension, and let \( A \) be a root system in \( V \). If \( \alpha, \beta \in A \), then let \( n(\alpha, \beta) \in Z \) be as in Section 19.9. Let \( (\cdot, \cdot) \) be an inner product on \( V \) that is invariant under the Weyl group of \( A \), and let \( \| \cdot \| \) be the corresponding norm on \( V \). If \( \alpha, \beta \in A \), then

\[ (20.15.1) \quad n(\alpha, \beta) = 2 (\alpha, \beta)/\|\beta\|^2, \]

as before. In particular, \( n(\alpha, \beta) = 0 \) if and only if \((\alpha, \beta) = 0\), which is symmetric in \( \alpha \) and \( \beta \). Otherwise, if \( n(\beta, \alpha) \neq 0 \), then

\[ (20.15.2) \quad n(\alpha, \beta)/\alpha(\beta, \alpha) = \|\alpha\|^2/\|\beta\|^2. \]

In this case, \( \|\alpha\| \) is greater than, equal to, or less than \( \|\beta\| \) exactly when (20.15.2) is greater than, equal to, or less than 1, respectively.

Let \( B \) be a base for \( A \). The corresponding Dynkin diagram adds some information to the Coxeter graph of \( A \) with respect to \( B \), so that the Cartan
matrix of $A$ with respect to $B$ is determined. More precisely, if $\alpha$ and $\beta$ are distinct elements of $B$, and if there is more than one edge between $\alpha$ and $\beta$ in the Coxeter graph, then $n(\alpha, \beta), n(\beta, \alpha) \neq 0$, and (20.15.2) is not equal to 1. In this situation, one can add a marking to the edges between $\alpha$ and $\beta$, to indicate whether (20.15.2) is greater than or less than 1. This is equivalent to indicating whether $\|\alpha\|$ is greater than or less than $\|\beta\|$, as on p57 of [13] and p39 of [23].

Alternatively, suppose that $A$ is irreducible, so that the Coxeter graph is connected. As on p38 of [23], one can label each $\alpha \in B$ with a positive real number that is proportional to $\|\alpha\|^2$. If $\alpha$ and $\beta$ are distinct elements of $B$, and if there is at least one edge between $\alpha$ and $\beta$ in the Coxeter graph, then $n(\alpha, \beta), n(\beta, \alpha) \neq 0$, and (20.15.2) is the same as the ratio of the labels of $\alpha$ and $\beta$. In this case, the number of edges between $\alpha$ and $\beta$ in the Coxeter graph is the same as this ratio. Two sets of labels of the elements of $B$ are considered to be equivalent when they differ by a single positive multiplicative constant.

Let $B_0$ be a nonempty subset of $B$, and let $V_0$ be the linear subspace spanned by $B_0$. Thus $A_0 = A \cap V_0$ is a root system in $V_0$, as in Section 20.6, and Exercise 2 on p54 of [13]. More precisely, $B_0$ is a base for $A_0$. If $\alpha \in A$, then let $\sigma_\alpha$ be the symmetry on $V$ with vector $\alpha$ that maps $A$ onto itself. If $\alpha \in A_0$, then $\sigma_\alpha$ maps $V_0$ onto itself, as before. This implies that $\sigma_\alpha$ maps $A_0$ onto itself, so that the restriction of $\sigma_\alpha$ to $V_0$ is the symmetry on $V_0$ with vector $\alpha$ that maps $A_0$ onto itself. If $\alpha, \beta \in A_0$, then it follows that the analogue of $n(\alpha, \beta)$ for $A_0$ is the same as for $A$. Similarly, the restriction of $\langle \cdot, \cdot \rangle$ to $V_0$ is invariant under the Weyl group of $A_0$. The Cartan matrix of $A_0$ with respect to $B_0$ is the same as the restriction to $B_0 \times B_0$ of the Cartan matrix of $A$ with respect to $B$. If $\alpha$ and $\beta$ are distinct elements of $B_0$, then the number of edges between $\alpha$ and $\beta$ in the Coxeter graph of $A_0$ with respect to $B_0$ is the same as in the Coxeter graph of $A$ with respect to $B$. One could also use the same additional markings or labellings for the Dynkin diagram of $A_0$ with respect to $B_0$ as for $A$ with respect to $B$. This corresponds to Exercise 6 on p54 of [13].
Chapter 21

Root systems, 3

21.1 More on type $A_2$

Let $(\cdot, \cdot)$ be the standard inner product on the complex plane, considered as a vector space over the real numbers, as in Section 20.13. Also let $A_{2,1}$ be the root system in $\mathbb{C}$ discussed in Section 20.14, consisting of the cube roots of unity and their negatives. Equivalently, $A_{2,1}$ consists of complex numbers of the form $\exp(j \pi i/3)$, with $j \in \mathbb{Z}$. Observe that

$$\exp(\pi i/3) + \exp(2 \pi i/3) = \sqrt{3} i,$$

by (20.14.1) and (20.14.2). If $j \in \mathbb{Z}$, then it follows that

$$\exp(j \pi i/3) + \exp((j + 1) \pi i/3) = \sqrt{3} i \exp((j - 1) \pi i/3).$$

Put

$$A_{2,2} = \sqrt{3} i A_{2,1} = \{ \sqrt{3} i z : z \in A_{2,1} \}.$$

This is the same as the set of complex numbers that can be expressed as in

(21.1.2) for some $j \in \mathbb{Z}$. It is clear from (21.1.3) that $A_{2,2}$ is a root system in $\mathbb{C}$, as a vector space over $\mathbb{R}$, that is isomorphic to $A_{2,1}$. If $w \in A_{2,2}$, then it is easy to see that

$$w/\bar{w} \in A_{2,1}.$$

Of course, this also holds when $w \in A_{2,1}$.

If $\gamma \in \mathbb{C} \setminus \{0\}$, then let $\sigma_\gamma$ be the reflection on $\mathbb{C}$, as a vector space over $\mathbb{R}$, associated to $\gamma$ with respect to $(\cdot, \cdot)$, as before. Thus $\sigma_\gamma(z) = -(\gamma/\bar{\gamma}) \bar{z}$ for every $z \in \mathbb{C}$, as in Section 20.13. If $\gamma \in A_{2,1}$ or $A_{2,2}$, then one can check that $\sigma_\gamma$ maps $A_{2,1}$ and $A_{2,2}$ onto themselves, using (21.1.4).

Remember that $\sigma_\gamma(z) - z = -2(z, \gamma) |\gamma|^{-2} \gamma$ for every $\gamma \in \mathbb{C} \setminus \{0\}$ and $z \in \mathbb{C}$, as in Section 20.13. Let us check that

$$2(z, \gamma) |\gamma|^{-2} = 2 \text{Re}(z/\gamma) \in \mathbb{Z}$$

(21.1.5)
for every $\gamma, z \in A_{2,1} \cup A_{2,2}$. This was already mentioned in Section 20.14 when $\gamma, z \in A_{2,1}$. If $\gamma, z \in A_{2,2}$, then this follows from the previous case, because $\gamma$ and $z$ are products of elements of $A_{2,1}$ by $\sqrt{3}i$. If $\gamma \in A_{2,1}$ and $z \in A_{2,2}$, then (21.1.5) can be reduced to the case where $z \in A_{2,1}$, because elements of $A_{2,2}$ can be expressed as sums of elements of $A_{2,1}$, as in (21.1.2). The case where $\gamma \in A_{2,2}$ and $z \in A_{2,1}$ is a more precise version of this, because of the additional factor of $|\gamma|^{-2} = 1/3$. If $\gamma = \sqrt{3}i$ and $z \in A_{2,1}$, then (21.1.5) can be verified directly. Otherwise, if $\gamma$ is the product of $\sqrt{3}i$ and any element of $A_{2,1}$, then one can reduce to the case where $\gamma = \sqrt{3}i$.

### 21.2 Type $G_2$

Some very nice pictures can be found on p44 of [13] and p27 of [23]. Let us continue with the notation in the previous section, and put

\[(21.2.1) \quad A = A_{2,1} \cup A_{2,2}.\]

It is easy to see that this defines a root system in the complex plane, as a vector space over the real numbers, using the remarks in the previous section. This root system is said to be of type $G_2$, and this description of it corresponds to the one on p40 of [23]. More precisely, $A$ is reduced as a root system in $\mathbb{C}$.

Put $\alpha = 1 \in A_{2,1}$ and

\[(21.2.2) \quad \beta = \sqrt{3}i \exp(\pi i/3) = \sqrt{3}i(1/2 + (\sqrt{3}/2)i) = -3/2 + (\sqrt{3}/2)i,\]

which is an element of $A_{2,2}$. Note that this choice of $\alpha$ is the same as in Section 20.14, but this choice of $\beta$ is different. Of course, $B = \{\alpha, \beta\}$ is a basis for $\mathbb{C}$ as a vector space over $\mathbb{R}$, and in fact it is a base for $A$, as in [13, 23]. Indeed,

\[(21.2.3) \quad \beta + \alpha = -1/2 + (\sqrt{3}/2)i = \exp(2\pi i/3)\]

and

\[(21.2.4) \quad \beta + 2\alpha = 1/2 + (\sqrt{3}/2)i = \exp(\pi i/3).\]

Similarly,

\[(21.2.5) \quad \beta + 3\alpha = 3/2 + (\sqrt{3}/2)i = \sqrt{3}i(1/2 - (\sqrt{3}/2)i) = \sqrt{3}i \exp(-\pi i/3).\]

We also have that

\[(21.2.6) \quad 2\beta + 3\alpha = \sqrt{3}i.\]

These expressions for elements of $A$ are indicated in the picture on p27 of [23], and the other elements of $A$ can be obtained by multiplying by $-1$.

If $z, w \in A$, then $n(z, w) = 2(z, w)|w|^{-2} \in \mathbb{Z}$, as in Section 19.9. Observe that

\[(21.2.7) \quad n(\alpha, \beta) = 2(\alpha, \beta)|\beta|^{-2} = 2 \Re(\alpha \overline{\beta})|\beta|^{-2} = -1\]
and
\[(21.2.8) \quad n(\beta, \alpha) = 2(\alpha, \beta) |\alpha|^{-2} = -3.\]

Thus the number of edges between \(\alpha\) and \(\beta\) in the Coxeter graph of \(A\) with respect to \(B\) is
\[(21.2.9) \quad n(\alpha, \beta) n(\beta, \alpha) = 3.\]

The corresponding Dynkin diagram may be depicted as on p38 of [23], where \(\alpha\) has the label \(|\alpha|^2 = 1\) and \(\beta\) has the label \(|\beta|^2 = 3\). Alternatively, the Dynkin diagram may be depicted as on p57f of [13] and p39 of [23], where the edges are marked to indicate that \(|\alpha| < |\beta|\).

Let \(T\) be a one-to-one linear mapping from \(C\) onto itself, as a vector space over \(R\). Remember that \(T\) is an automorphism of the root system \(A\) if \(T(A) = A\). In this case, if \(T\) maps \(B\) onto itself, then \(T\) preserves the Cartan matrix of \(A\) with respect to \(B\), as in Section 20.2. It is easy to see that this only happens when \(T\) is the identity mapping on \(C\) in this situation. This is mentioned in Table 1 on p66 of [13], and on p39 of [23].

If \(z \in C \setminus \{0\}\), then put \(\widehat{\mathbb{F}} = \{z/|z|^2\}\), as in (19.7.9). If \(E \subseteq C \setminus \{0\}\), then let \(\widehat{E}\) be the set of \(\widehat{z}, z \in E\). Thus
\[(21.2.10) \quad \widehat{A}_{2,1} = 2A_{2,1} = \{2z : z \in A_{2,1}\}\]
and
\[(21.2.11) \quad \widehat{A}_{2,2} = 2/3A_{2,2} = \{(2/3)z : z \in A_{2,2}\}.\]

Observe that
\[(21.2.12) \quad \widehat{A}_{2,1} = (2/\sqrt{3}) i A_{2,2}\]
and
\[(21.2.13) \quad \widehat{A}_{2,2} = (2/\sqrt{3}) i A_{2,1}.\]

This implies that
\[(21.2.14) \quad \widehat{A} = \widehat{A}_{2,1} \cup \widehat{A}_{2,2} = (2/\sqrt{3}) i A.\]

Remember that \(\widehat{A}\) is the inverse system of \(A\) in \(C\) with respect to \((\cdot, \cdot)\), as in Section 19.8. It follows that \(\widehat{A}\) is isomorphic to \(A\), as root systems in \(C\). This corresponds to part of Exercise 5 on p63 of [13].

### 21.3 Another description

Let us now consider the description on p65 of [13], of a root system that is isomorphic to the one in the previous section. Let \((\cdot, \cdot)\) be the standard inner product on \(R^3\), and let \(e_1, e_2,\) and \(e_3\) be the standard basis vectors in \(R^3\). Put
\[(21.3.1) \quad V = \{v \in R^3 : v_1 + v_2 + v_3 = 0\},\]
which is the two-dimensional linear subspace of \(R^3\) orthogonal to \(e_1 + e_2 + e_3\).

Consider
\[(21.3.2) \quad A_1 = \{\alpha \in V \cap Z^3 : (\alpha, \alpha) = 2\},\]
\[(21.3.3) \quad A_2 = \{\alpha \in V \cap Z^3 : (\alpha, \alpha) = 6\},\]
432  CHAPTER 21.  ROOT SYSTEMS, 3

and

\[(21.3.4) \quad A = A_1 \cup A_2.\]

Thus $V$ is the same as in Section 20.10 with $n = 2$, and $A_1$ is the root system discussed there in this case. As before, the elements of $A_1$ are of the form $e_j - e_l$, where $1 \leq j \neq l \leq 3$. Observe that $\alpha \in \mathbb{Z}^3$ satisfies $(\alpha, \alpha) = 6$ exactly when two of the coordinates of $\alpha$ are $\pm 1$, and the third coordinate is $\pm 2$. It follows that $\alpha \in A_2$ exactly when one of the coordinates of $\alpha$ is $2$ and the other two coordinates are $-1$, or one of the coordinates is $-2$ and the other two coordinates are $1$. In particular, $A$ is a finite set of nonzero elements of $V$ whose linear span is $V$.

If $\alpha$ is a nonzero element of $\mathbb{R}^3$, then let $\sigma_\alpha$ be the reflection on $\mathbb{R}^3$ associated to $\alpha$ with respect to $(\cdot, \cdot)$, as usual. If $\alpha \in V$, then

\[(21.3.5) \quad \sigma_\alpha(V) = V,\]

as in Section 20.6. Suppose for the moment that $(\alpha, \alpha) = 2$, so that $\sigma_\alpha(v) = v - (v, \alpha) \alpha$ for every $v \in \mathbb{R}^3$. If $\alpha \in \mathbb{Z}^3$, then $\sigma_\alpha(\mathbb{Z}^3) = \mathbb{Z}^3$, as in Section 20.10. If $\alpha \in V \cap \mathbb{Z}^3$, then it follows that $\sigma_\alpha$ maps $A_1$ and $A_2$ onto themselves, because $\sigma_\alpha$ is an orthogonal transformation on $\mathbb{R}^3$. Of course, this implies that $\sigma_\alpha$ maps $A$ onto itself. If $\alpha, v \in \mathbb{Z}^3$, then $(v, \alpha) \in \mathbb{Z}$, so that $\sigma_\alpha(v) - v = (v, \alpha) \alpha$ is an integer multiple of $\alpha$. If $\alpha = e_j - e_l$, $1 \leq j \neq l \leq 3$, then $\sigma_\alpha(v)$ interchanges the $j$th and $l$th coordinates of $v$, without affecting the other coordinate, as before.

Suppose now that $\alpha \in A_2$, so that

\[(21.3.6) \quad \sigma_\alpha(v) = v - (v, \alpha) \alpha/3\]

for every $v \in \mathbb{R}^3$. If $\beta \in A$, then we would like to check that

\[(21.3.7) \quad (\beta, \alpha)/3 \in \mathbb{Z}.\]

Suppose first that $\beta \in A_1$, so that $\beta = e_j - e_l$ for some $1 \leq j \neq l \leq n$. If $\alpha_j = \alpha_l$, then $(\beta, \alpha) = 0$. Otherwise, if $\alpha_j \neq \alpha_l$, then one can check that

\[(21.3.8) \quad (\beta, \alpha) = \pm 3.\]

If $\beta = \pm \alpha \in A_2$, then (21.3.7) clearly holds. If $\beta \in A_2$ and $\beta \neq \pm \alpha$, then one can verify that (21.3.8) holds.

It follows from (21.3.7) that $\sigma_\alpha$ maps $A$ into $\mathbb{Z}^3$. This implies that $\sigma_\alpha$ maps $A_1$ and $A_2$ into themselves, because $\sigma_\alpha$ is an orthogonal transformation on $\mathbb{R}^3$ that maps $V$ to itself. This means that $\sigma_\alpha$ maps $A_1$ and $A_2$ onto themselves, because $\sigma_\alpha$ is its own inverse, so that $\sigma_\alpha$ maps $A$ onto itself as well. If $\beta \in A$, then $\sigma_\alpha(\beta) - \beta$ is an integer multiple of $\alpha$, by (21.3.7). This shows that $A$ is a root system in $V$, and it is easy to see that $A$ is reduced.

As a base for $A$, one can take

\[(21.3.9) \quad \{e_1 - e_2, -2e_1 + e_2 + e_3\},\]
21.4. TYPE $F_4$

In this section, we take $V = \mathbb{R}^4$, with its standard inner product $(\cdot, \cdot)$ and basis $e_1, e_2, e_3, e_4$. Put

$$A_1 = \{ \alpha \in \mathbb{Z}^4 : (\alpha, \alpha) = 1 \text{ or } 2 \},$$

as in Section 20.11, with $n = 4$. The elements of $A_1$ with $(\alpha, \alpha) = 1$ are of the form $\pm e_j$ for some $j = 1, 2, 3, 4$, and the elements of $A_1$ with $(\alpha, \alpha) = 2$ are of the form $\pm e_j \pm e_l$ for some $1 \leq j \neq l \leq n$, as before. We have seen that $A_1$ is a reduced root system in $V$.

Let $\tilde{L}_4$ be the subgroup of $\mathbb{R}^4$, as a commutative group with respect to addition, generated by $e_1, e_2, e_3, e_4$, and $(1/2) (e_1 + e_2 + e_3 + e_4)$. Of course, $\tilde{L}_4 = \mathbb{Z}^4$ is the subgroup of $\mathbb{R}^4$ generated by $e_1, e_2, e_3, e_4$, so that $\tilde{L}_4$ is the same as the subgroup of $\mathbb{R}^4$ generated by $\mathbb{Z}^4$ and $(1/2) (e_1 + e_2 + e_3 + e_4)$. Equivalently,

$$\tilde{L}_4 = \mathbb{Z}^4 \cup (\mathbb{Z}^4 + (1/2, 1/2, 1/2, 1/2)).$$

Consider

$$A = \{ \alpha \in \tilde{L}_4 : (\alpha, \alpha) = 1 \text{ or } 2 \}.$$
Note that
\[(21.4.4) \quad A \cap \mathbb{Z}^n = A_1.\]
One can check that
\[(21.4.5) \quad A \cap (\mathbb{Z}^4 + (1/2, 1/2, 1/2, 1/2)) = \{\alpha \in \mathbb{R}^4 : \alpha_j = \pm 1/2 \text{ for every } j = 1, 2, 3, 4\}.\]

More precisely, it is easy to see that if \(\alpha_j = \pm 1/2\) for each \(j = 1, 2, 3, 4\), then \(\alpha \in \mathbb{L}_4\) and \((\alpha, \alpha) = 1\), so that \(\alpha \in A\). If \(\alpha \in \mathbb{Z}^4 + (1/2, 1/2, 1/2, 1/2)\) and \(\alpha_j \neq \pm 1/2\) for some \(j = 1, 2, 3, 4\), then one can verify that \((\alpha, \alpha) \geq 3\).

Of course, \(A\) is a finite set of nonzero elements of \(V\), whose linear span is \(V\). If \(\alpha \in V \setminus \{0\}\), then let \(\sigma_\alpha\) be the reflection on \(V\) associated to \(\alpha\) with respect to \((\cdot, \cdot)\), as usual. Suppose first that \(\alpha = \pm e_j\) for some \(j = 1, 2, 3, 4\), so that \(\sigma_\alpha = \sigma_{e_j}\) multiplies the \(j\)th coordinate of \(v \in V\) by \(-1\), and leaves the other coordinates of \(v\) unchanged. Clearly \(\sigma_\alpha\) maps \(\mathbb{Z}^4, \mathbb{L}_4, A_1,\) and \(A\) onto themselves. If \(v \in \mathbb{L}_4\), then \(2v_j \in \mathbb{Z}\), and \(\sigma_\alpha(v) - v = -2v_je_j\) is an integer multiple of \(\alpha\).

Suppose now that \(\alpha = \pm e_j \pm e_l\) for some \(1 \leq j \neq l \leq 4\), so that \(\sigma_\alpha(v) = v - (v, \alpha)\alpha\) for every \(v \in V\). More precisely, we have seen that \(\sigma_\alpha\) interchanges the \(j\)th and \(l\)th coordinates of \(v \in V\) in this case, and may multiply them by \(-1\), while leaving the other coordinates of \(v\) unchanged. In particular, \(\sigma_{e_j}\) maps \(\mathbb{Z}^4, \mathbb{L}_4, A_1,\) and \(A\) onto themselves. If \(v \in \mathbb{L}_4\), then one can check that
\[(21.4.6) \quad (v, \alpha) \in \mathbb{Z}.\]

This implies that \(\sigma_\alpha(v) - v = -(v, \alpha)\alpha\) is an integer multiple of \(\alpha\) when \(v \in \mathbb{L}_4\), and in particular when \(v \in A\).

Suppose that \(\alpha\) is an element of \((21.4.5)\), so that \(\alpha_j = \pm 1/2\) for each \(j = 1, 2, 3, 4\). If \(v \in \mathbb{L}_4\), then one can verify that
\[(21.4.7) \quad 2(v, \alpha) \in \mathbb{Z}.\]

More precisely, if \(v \in \mathbb{Z}^4\), then this follows from the fact that \(2\alpha \in \mathbb{Z}^4\). If \(v = (1/2, 1/2, 1/2, 1/2)\), then \((21.4.7)\) can be checked directly. This implies that \((21.4.7)\) holds for all \(v \in \mathbb{L}_4\).

Note that \(\sigma_\alpha(v) = v - 2(v, \alpha)\alpha\) for every \(v \in V\), because \((\alpha, \alpha) = 1\). It follows that \(\sigma_\alpha\) maps \(\mathbb{L}_4\) into itself, by \((21.4.7)\). Thus \(\sigma_\alpha\) maps \(\mathbb{L}_4\) onto itself, because \(\sigma_\alpha\) is its own inverse on \(V\). This means that \(\sigma_\alpha\) maps \(A\) onto itself, because \(\sigma_\alpha\) is an orthogonal transformation on \(V\). If \(v \in \mathbb{L}_4\), then \(\sigma_\alpha(v) - v\) is an integer multiple of \(\alpha\), by \((21.4.7)\).

This shows that \(A\) is a root system in \(V\), which is said to be of type \(F_4\). This follows the descriptions on p65 of [13], and p40 of [23]. Observe that \(A\) is reduced, as a root system in \(V\).

Put
\[(21.4.8) \quad V_0 = \{v \in V : v_1 = 0\}\]
and
\[(21.4.9) \quad A_0 = A \cap V_0 = A_1 \cap V_0.\]

Observe that $A_0$ is a reduced root system in $V_0$, which is isomorphic to the analogue of $A_1$ in $\mathbb{R}^3$.

Put
\[(21.4.10) \quad B = \{e_2 - e_3, e_3 - e_4, e_4, (1/2)(e_1 - e_2 - e_3 - e_4)\}\]
and
\[(21.4.11) \quad B_0 = B \cap V_0 = \{e_2 - e_3, e_3 - e_4, e_4\}.\]

It is easy to see that $B_0$ is a basis for $V_0$. In fact, $B_0$ is a base for $A_0$ as a root system in $V_0$, as in Section 20.11. We also have that $B$ is a basis for $V$, because $B_0$ is a basis for $V_0$. One can check that $B$ is a base for $A$, as in [13, 23].

More precisely, note that $e_2$ and $e_3$ can be expressed as sums of elements of $B_0$, as in Section 20.11. Using this, one can express any element of $A$ as
\[e_1 - e_2 - e_3 - e_4 \quad \text{and} \quad e_1 + e_2, \quad e_1, \quad e_1 + e_1, \quad \text{as sums of (21.4.12) and elements of } B_0.\]

The maximal root in $A$ with respect to $B$ is
\[(21.4.13) \quad 2(e_2 - e_3) + 3(e_3 - e_4) + 4e_4 + 2((1/2)(e_1 - e_2 - e_3 - e_4)) = e_1 + e_2,\]
as in Table 2 on p66 of [13]. The maximal short root is
\[(21.4.14) \quad (e_2 - e_3) + 2(e_3 - e_4) + 3e_4 + 2((1/2)(e_1 - e_2 - e_3 - e_4)) = e_1.\]

If $\alpha, \beta \in A$, then $n(\alpha, \beta) = 2(\alpha, \beta)(\beta, \beta)^{-1} \in \mathbb{Z}$, as in Section 19.9. If $\alpha, \beta \in B$ and $\alpha \neq \beta$, then the number of edges between $\alpha$ and $\beta$ in the Coxeter graph of $A$ with respect to $B$ is $n(\alpha, \beta) n(\beta, \alpha)$. In particular, if $(\alpha, \beta) = 0$, then $n(\alpha, \beta) = n(\beta, \alpha) = 0$, and there are no edges between $\alpha$ and $\beta$ in the Coxeter graph of $A$ with respect to $B$. Clearly
\[(21.4.15) \quad (e_2 - e_3, e_4) = (e_2 - e_3, (1/2)(e_1 - e_2 - e_3 - e_4)) = (e_3 - e_4, (1/2)(e_1 - e_2 - e_3 - e_4)) = 0.\]

Thus there are no edges between $e_2 - e_3$ and $e_4$, or between either $e_2 - e_3$ or $e_3 - e_4$ and $(1/2)(e_1 - e_2 - e_3 - e_4)$ in the Coxeter graph of $A$ with respect to $B$.

It is easy to see that
\[(21.4.16) \quad n(e_2 - e_3, e_3 - e_4) = n(e_3 - e_4, e_2 - e_3) = -1.\]
and

\[(21.4.17)\]
\[n(e_4, (1/2) (e_1 - e_2 - e_3 - e_4)) = n((1/2) (e_1 - e_2 - e_3 - e_4), e_4) = -1.\]

This implies that there is one edge between $e_2 - e_3$ and $e_3 - e_4$, and one edge between $e_4$ and $(1/2) (e_1 - e_2 - e_3 - e_4)$, in the Coxeter graph of $A$ with respect to $B$.

We also have that
\[(21.4.18)\]
\[n(e_3 - e_4, e_4) = -2\]
and
\[(21.4.19)\]
\[n(e_4 - e_3, e_4) = -1.\]

This means that there are two edges between $e_3 - e_4$ and $e_4$ in the Coxeter graph of $A$ with respect to $B$. The corresponding Dynkin diagram may be depicted as on p38 of [23], where $e_2 - e_3$ and $e_3 - e_4$ are labelled with 2, and $e_4$ and $(1/2) (e_1 - e_2 - e_3 - e_4)$ are labelled with 1. The Dynkin diagram may be depicted as on p57f of [13] and p39 of [23] as well, where the edges between $e_3 - e_4$ and $e_4$ are marked to indicate that $(e_3 - e_4, e_3 - e_4) = 2 > 1 = (e_4, e_4)$.

Let $T$ be an automorphism of $A$ as a root system, so that $T$ is a one-to-one linear mapping from $V$ onto itself that maps $A$ onto itself. If $T(B) = B$, then $T$ preserves the Cartan matrix of $A$ with respect to $B$, as in Section 20.2. It is easy to see that this can only happen when $T$ is the identity mapping on $V$, as indicated in Table 1 on p66 of [13], and on p39 of [23].

### 21.5 Inverse systems

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $A$ be a root system in $V$. If $\alpha, \beta \in A$, then we let $n(\alpha, \beta) \in \mathbb{Z}$ be as in Section 19.9, as usual. Let $(\cdot, \cdot)$ be an inner product on $V$ that is invariant under the Weyl group of $A$, so that $n(\alpha, \beta) = 2 (\alpha, \beta) (\beta, \beta)^{-1}$ for every $\alpha, \beta \in A$. If $v \in V \setminus \{0\}$, then put $\tilde{v} = 2v (v, v)^{-1}$, as in (19.7.9). Note that

\[(21.5.1)\]
\[(\tilde{v}, \tilde{v}) = 4/(v, v).\]

Remember that $\tilde{A} = \{ \tilde{\alpha} : \alpha \in A \}$ is the inverse system of $A$ in $V$ with respect to $(\cdot, \cdot)$, as in Section 19.8. If $\alpha, \beta \in A$, then $\tilde{\alpha}, \tilde{\beta} \in \tilde{A}$, and $n(\tilde{\alpha}, \tilde{\beta})$ can be defined as before. Observe that

\[(21.5.2)\]
\[n(\tilde{\alpha}, \tilde{\beta}) = 2(\tilde{\alpha}, \tilde{\beta})(\beta, \beta)^{-1} = 2(\alpha, \beta)(\alpha, \alpha)^{-1} = n(\beta, \alpha).\]

It follows that

\[(21.5.3)\]
\[n(\tilde{\alpha}, \tilde{\beta}) n(\tilde{\beta}, \tilde{\alpha}) = n(\alpha, \beta) n(\beta, \alpha)\]

for every $\alpha, \beta \in A$. Of course, if $(\alpha, \alpha) = (\beta, \beta)$, then

\[(21.5.4)\]
\[n(\tilde{\alpha}, \tilde{\beta}) = n(\alpha, \beta).\]
21.6. TYPE $E_8$

Let $B$ be a base for $A$. Suppose for the moment that $A$ is reduced, so that $\hat{B} = \{\hat{\alpha} : \alpha \in B\}$ is a base for $\hat{A}$, as in Section 19.13. In this case, the Coxeter graph of $A$ with respect to $B$ corresponds to the Coxeter graph of $\hat{A}$ with respect to $\hat{B}$ under the mapping $\alpha \mapsto \hat{\alpha}$, by (21.5.3). If $A$ is not reduced, then one can get a base for $\hat{A}$ from $B$ as in Section 19.13. One can check that the Coxeter graphs for $A$ and $\hat{A}$ with respect to these bases correspond to each other in this situation as well.

If the elements of $A$ have the same norm with respect to $(\cdot, \cdot)$, then there is a dilation on $V$ that sends $A$ onto $\hat{A}$. This dilation sends $\alpha \in A$ to $\hat{\alpha} \in \hat{A}$.

If $A$ is as in Section 21.2, then the elements of $A$ do not all have the same norm, but we saw that $A$ is isomorphic to $A$. Alternatively, let $B = \{\alpha, \beta\}$ be a base for $A$, so that $\hat{B} = \{\hat{\alpha}, \hat{\beta}\}$ is a base for $\hat{A}$. Consider the mapping $\phi$ from $B$ onto $\hat{B}$ defined by $\phi(\alpha) = \hat{\beta}$ and $\phi(\beta) = \hat{\alpha}$. Using $\phi$, the Cartan matrix of $A$ with respect to $B$ corresponds to the Cartan matrix of $\hat{A}$ with respect to $\hat{B}$, because of (21.5.1). Thus $\phi$ extends to an isomorphism between $A$ and $\hat{A}$, as in Section 20.2.

Let us now return to the situation considered in the previous section. Note that $\hat{\alpha} = 2 \alpha$ when $\alpha \in V$ satisfies $(\alpha, \alpha) = 1$, and $\hat{\alpha} = \alpha$ when $(\alpha, \alpha) = 2$. If $A_1$ is as in (21.4.1), then its inverse system $\hat{A}_1$ in $V = \mathbb{R}^4$ with respect to the standard inner product consists of elements of the form $\pm 2e_j$ for $j = 1, 2, 3, 4$ and $\pm e_j \pm e_l$ for $1 \leq j \neq l \leq 4$, as in Section 20.11. If $A$ is as in (21.4.3), then the inverse system $\hat{A}$ of $A$ in $V$ with respect to the standard inner product consists of the elements of $\hat{A}_1$, and elements of the form $\pm e_1 \pm e_2 \pm e_3 \pm e_4$.

Let $B$ be the base for $A$ in (21.4.10). The corresponding base $\hat{B}$ for $\hat{A}$ is given by

$$\hat{B} = \{e_2 - e_3, e_3 - e_4, 2e_4, e_1 - e_2 - e_3 - e_4\}.$$ (21.5.5)

Let $\phi$ be the one-to-one mapping from $B$ onto $\hat{B}$ that interchanges the ordering in these lists. One can verify that the Cartan matrix of $A$ with respect to $B$ corresponds to the Cartan matrix of $\hat{A}$ with respect to $\hat{B}$, using $\phi$. This means that $\phi$ extends to an isomorphism between $A$ and $\hat{A}$, as in Section 20.2 again.

This corresponds to part of Exercise 5 on p63 of [13]. Of course, the Dynkin diagram of the inverse system can be obtained by changing the labels on the vertices, or the markings of the multiple edges. In the previous two cases, the resulting Dynkin diagram is easily seen to be isomorphic to the initial one.

21.6 Type $E_8$

Let us take $V = \mathbb{R}^8$, with its standard inner product $(\cdot, \cdot)$ and basis $e_1, \ldots, e_8$.

Put

$$(21.6.1) \quad A_1 = \{\alpha \in \mathbb{Z}^8 : (\alpha, \alpha) = 2\},$$

as in Section 20.12, with $n = 8$. The elements of $A_1$ are of the form $\pm e_j \pm e_l$ with $1 \leq j \neq l \leq 8$, as before. Remember that $A$ is a reduced root system in $V$. 

Of course, \( L_8 = \mathbb{Z}^8 \) is the subgroup of \( \mathbb{R}^8 \), as a commutative group with respect to addition, generated by \( e_1, \ldots, e_8 \). Let \( \tilde{L}_8 \) be the subgroup of \( \mathbb{R}^8 \) generated by \( e_1, \ldots, e_8 \) and \((1/2)(e_1 + \cdots + e_8)\). Thus

\[
\tilde{L}_8 = \mathbb{Z}^8 \cup (\mathbb{Z}^8 + (1/2, \ldots, 1/2)).
\]

If \( \alpha \in \mathbb{R}^8 \) satisfies \( \alpha_j = \pm 1/2 \) for every \( j = 1, \ldots, 8 \), then \( \alpha \) is an element of \( \mathbb{Z}^8 + (1/2, \ldots, 1/2) \) and \( (\alpha, \alpha) = 2 \). If \( \alpha \in \mathbb{Z}^8 + (1/2, \ldots, 1/2) \) and \( \alpha_j \neq \pm 1/2 \) for some \( j = 1, \ldots, 8 \), then it is easy to see that \( (\alpha, \alpha) \geq 4 \).

Let \( L_{8,e} \) be the subgroup of \( \mathbb{Z}^8 \) such that

\[
\sum_{j=1}^{8} v_j \in 2\mathbb{Z},
\]

and let \( \tilde{L}_{8,e} \) be the set of \( v \in \tilde{L}_8 \) such that (21.6.3) holds. These are subgroups of \( \mathbb{Z}^8 \) and \( \tilde{L}_8 \), respectively. Note that

\[
(1/2)(e_1 + \cdots + e_8) \in \tilde{L}_{8,e},
\]

because the sum of its coordinates is 4. It follows that

\[
\tilde{L}_{8,e} = L_{8,e} \cup (L_{8,e} + (1/2, \ldots, 1/2)).
\]

We also have that \( A_1 \subseteq L_{8,e} \).

Suppose that \( \alpha \in \mathbb{R}^8 \) satisfies \( \alpha_j = \pm 1/2 \) for each \( j = 1, \ldots, 8 \), and let \( r \) be the number of \( j \) such that \( \alpha_j = 1/2 \). In this case,

\[
\sum_{j=1}^{8} \alpha_j = (1/2)(r - (8 - r)) = r - 4.
\]

This means that \( \alpha \in \tilde{L}_{8,e} \) exactly when \( r \) is even.

Consider

\[
A = \{ \alpha \in \tilde{L}_{8,e} : (\alpha, \alpha) = 2 \}.
\]

Observe that

\[
A \cap L_{8,e} = A_1.
\]

Using the previous remarks, we get that \( \alpha \in A \cap (L_{8,e} + (1/2, \ldots, 1/2)) \) if and only if \( \alpha_j = \pm 1/2 \) for every \( j = 1, \ldots, 8 \), and \( \alpha_j = 1/2 \) for an even number of \( j \). Clearly \( A \) is a finite set of nonzero elements of \( V \) whose linear span is \( V \).

Let us check that

\[
(v, w) \in \mathbb{Z}
\]

for every \( v, w \in \tilde{L}_{8,e} \). Of course, (21.6.9) holds when \( v, w \in \mathbb{Z}^8 \). If \( w = (1/2, \ldots, 1/2) \), then (21.6.9) holds for every \( v \in L_{8,e} \), by (21.6.3). If \( v = (1/2, \ldots, 1/2) \) as well, then \( (v, w) = 2 \). It follows that (21.6.9) holds for every \( v, w \in \tilde{L}_{8,e} \), by (21.6.5).
21.6. TYPE $E_8$

If $\alpha \in V$ satisfies $(\alpha, \alpha) = 2$, then the reflection $\sigma_\alpha$ on $V$ associated to $\alpha$ with respect to $(\cdot, \cdot)$ is given by $\sigma_\alpha(v) = v - (v, \alpha)\alpha$, as before. If $\alpha \in A$, then $\sigma_\alpha$ maps $\tilde{L}_{8,e}$ into itself, because of (21.6.9). This implies that $\sigma_\alpha$ maps $L_{8,e}$ onto itself, because $\sigma_\alpha$ is its own inverse. It follows that $\sigma_\alpha$ maps $A$ onto itself, because $\sigma_\alpha$ is an orthogonal transformation on $V$. If $v \in \tilde{L}_{8,e}$, then $\sigma_\alpha(v) - v$ is an integer multiple of $\alpha$, by (21.6.9).

Thus $A$ is a root system in $V$, which is said to be of type $E_8$. This follows the descriptions on p65 of [13], and p41 of [23]. Note that $A$ is reduced, as a root system in $V$.

Put

\[(21.6.10) \quad V_0 = \{v \in V : e_8 = 0\}\]

and

\[(21.6.11) \quad A_0 = A \cap V_0.\]

Equivalently,

\[(21.6.12) \quad \tilde{L}_{8,e} \cap V_0 = L_{8,e} \cap V_0.\]

because $\tilde{L}_{8,e} \cap V_0 = L_{8,e} \cap V_0$. This is a reduced root system in $V_0$, which is isomorphic to the one in Section 20.12 with $n = 7$.

Put

\[(21.6.13) \quad B = \{(1/2)(e_1 + e_8 - (e_2 + \cdots + e_7)), e_1 + e_2, e_2 - e_1, e_3 - e_2, e_4 - e_3, e_5 - e_4, e_6 - e_5, e_7 - e_6\}\]

and

\[(21.6.14) \quad B_0 = B \cap V_0 = \{e_1 + e_2, e_2 - e_1, e_3 - e_2, e_4 - e_3, e_5 - e_4, e_6 - e_5, e_7 - e_6\}.\]

Observe that $B_0$ is a basis for $V_0$. More precisely, $B_0$ is a base for $A_0$, as a root system in $V_0$. This is similar to the base discussed in Section 20.12, with $n = 7$, ordering the standard basis vectors in $V_0$ the other way. If $1 \leq l < j \leq 7$, then $e_j - e_l$ and $e_j + e_l$ can be expressed as sums of elements of $B_0$, as before.

It is easy to see that $B$ is a basis for $V$, because $B_0$ is a basis for $V_0$. In fact, $B$ is a base for $A$, as on p65 of [13]. More precisely, suppose first that $\alpha \in A$ satisfies $\alpha_j = \pm 1/2$ for $j = 1, \ldots, 7$, and $\alpha_8 = 1/2$. One can express $\alpha$ as $(1/2)(e_1 + e_8 - (e_2 + \cdots + e_7))$ plus possibly some elements elements of $B_0$.

One can express $e_8 - e_7$ as the sum of

\[(21.6.15) \quad 2((1/2)(e_1 + e_8 - (e_2 + \cdots + e_7))) = e_8 + e_1 - (e_2 + \cdots + e_7)\]

and elements of $B_0$. Using this, one can express $e_8 - e_l$ as the sum of (21.6.15) and elements of $B_0$ for every $l = 1, \ldots, 7$. One can use this to express $e_8 + e_l$ as the sum of (21.6.15) and elements of $B_0$ for every $l = 1, \ldots, 7$.

The maximal root in $A$ with respect to $B$ is

\[(21.6.16) \quad 2((1/2)(e_1 + e_8 - (e_2 + \cdots + e_7))) + 3(e_1 + e_2) + 4(e_2 - e_1) + 6(e_3 - e_2) + 5(e_4 - e_3) + 4(e_5 - e_4) + 3(e_6 - e_5) + 2(e_7 - e_6) = e_8 + e_7,\]
as in Table 2 on p66 of [13].

If $\alpha, \beta \in A$, then let $n(\alpha, \beta) \in \mathbb{Z}$ be as in Section 19.9, so that $n(\alpha, \beta) = 2(\alpha, \beta)(\beta, \beta)^{-1} = (\alpha, \beta)$. If $\alpha, \beta \in B$ and $\alpha \neq \beta$, then $n(\alpha, \beta)n(\beta, \alpha) = (\alpha, \beta)^2$ is the number of edges between $\alpha$ and $\beta$ in the Coxeter graph of $A$ with respect to $B$. Observe that

$$\text{(21.6.17)} \quad (1/2)(e_1 + e_8 - (e_2 + \cdots + e_7)), e_2 - e_1) = -1.$$  

However, $(1/2)(e_1 + e_8 - (e_2 + \cdots + e_7))$ is orthogonal to the elements of $B$ other than $e_2 - e_1$ and itself.

The inner products between distinct elements of $B_0$ can be determined as in Section 20.12. More precisely,

$$\text{(21.6.18)} \quad (e_1 + e_2, e_3 - e_2) = -1,$$

and $e_1 + e_2$ is orthogonal to the elements of $B_0$ other than $e_3 - e_2$ and itself. Similarly,

$$\text{(21.6.19)} \quad (e_j - e_{j-1}, e_l - e_{l-1})$$

is equal to $-1$ when $|j - l| = 1$, and to $0$ when $|j - l| \geq 2$. The Coxeter graph or Dynkin diagram of $A$ can be depicted as on p58 of [13], or p37, 39 of [23].

Let $T$ be an automorphism of $A$ as a root system in $V$, and suppose that $T(B) = B$. This implies that $T$ preserves the Cartan matrix of $A$ with respect to $B$, as in Section 20.2, so that $T$ preserves the Coxeter graph of $A$ with respect to $B$ too. One can check that this can only happen when $T$ is the identity mapping on $V$, as indicated in Table 1 on p66 of [13], and on p39 of [23].

### 21.7 Type $E_7$

Let us continue with the same notation and hypotheses as in the previous section. Observe that

$$\text{(21.7.1)} \quad (1/2)(e_1 + e_8 - (e_2 + \cdots + e_7)) \in \mathbb{Z}^8 + (1/2, \ldots, 1/2) \subseteq \tilde{L}_8,$$

so that

$$\text{(21.7.2)} \quad \mathbb{Z}^8 + (1/2, -1/2, \ldots, -1/2, 1/2) = \mathbb{Z}^8 + (1/2, \ldots, 1/2).$$

Thus

$$\text{(21.7.3)} \quad \tilde{L}_8 = \mathbb{Z}^8 \cup (\mathbb{Z}^8 + (1/2, -1/2, \ldots, -1/2, 1/2)),$$

which is the same as the subgroup of $\mathbb{R}^8$, as a commutative group with respect to addition, generated by $e_1, \ldots, e_8$ and $(1/2)(e_1 + e_8 - (e_2 + \cdots + e_7))$. More precisely,

$$\text{(21.7.4)} \quad (1/2)(e_1 + e_8 - (e_2 + \cdots + e_7)) \in \tilde{L}_8, e,$$

with the sum of its coordinates equal to $-2$, and in fact

$$\text{(21.7.5)} \quad (1/2)(e_1 + e_8 - (e_2 + \cdots + e_7)) \in L_{8,e} + (1/2, \ldots, 1/2).$$
21.7. TYPE $E_7$

This implies that

$$L_{8,e} + (1/2, -1/2, \ldots, -1/2, 1/2) = L_{8,e} + (1/2, \ldots, 1/2),$$

so that

$$\tilde{L}_{8,e} = L_{8,e} \cup (L_{8,e} + (1/2, -1/2, \ldots, -1/2, 1/2)).$$

Remember that $V = \mathbb{R}^8$, and put

$$V_7 = \{v \in V : v_7 = -v_8\}.$$  

(21.7.6)

Note that $(1/2) (e_1 + e_8 - (e_2 + \cdots + e_7)) \in V_7$. The elements of

$$A_1 \cap V_7 = \{\alpha \in \mathbb{Z}^8 \cap V_7 : (\alpha, \alpha) = 2\}$$

(21.7.7)

are of the form $\pm e_j \pm e_l$ with $1 \leq j \neq l \leq 6$, or $\pm (e_7 - e_8)$. It is easy to see that $A_1 \cap V_7$ spans $V_7$. In fact, $A_1 \cap V_7$ is a reduced root system in $V_7$, as in Section 20.6.

Consider

$$A_7 = A \cap V_7 = \{\alpha \in \tilde{L}_{8,e} \cap V_7 : (\alpha, \alpha) = 2\}.$$  

(21.7.8)

Clearly

$$A_7 \cap L_{8,e} = (A \cap L_{8,e}) \cap V_7 = A_1 \cap V_7.$$  

(21.7.9)

Observe that

$$A_7 \cap (L_{8,e} + (1/2, -1/2, \ldots, -1/2, 1/2)) = (A \cap (L_{8,e} + (1/2, \ldots, 1/2))) \cap V_7$$

(21.7.10)

consists of $\alpha \in \mathbb{R}^8$ such that $\alpha_j = \pm 1/2$ for every $j = 1, \ldots, 8$, $\alpha_j = 1/2$ for an even number of $j$, and $\alpha_7 = -\alpha_8$. Of course, $V_7$ is spanned by $A_7$, because $V_7$ is spanned by $A_1 \cap V_7$. Thus $A_7$ is a root system in $V_7$, as in Section 20.6. This root system is said to be of type $E_7$. Note that $A_7$ is reduced, as a root system in $V_7$.

Put

$$V_{7,0} = V_7 \cap V_0 = \{v \in V : v_7 = v_8 = 0\}$$

(21.7.11)

and

$$A_{7,0} = A_7 \cap V_{7,0} = A \cap V_{7,0}.$$  

(21.7.12)

Equivalently,

$$A_{7,0} = A_1 \cap V_{7,0},$$

(21.7.13)

because $\tilde{L}_{8,e} \cap V_{7,0} = L_{8,e} \cap V_{7,0}$. This is a reduced root system in $V_{7,0}$, which is isomorphic to the one in Section 20.12 with $n = 6$.

Put

$$B_7 = \{(1/2) (e_1 + e_8 - (e_2 + \cdots + e_7)), e_1 + e_2, e_2 - e_1, e_3 - e_2, e_4 - e_3, e_5 - e_4, e_6 - e_5\}.$$  

(21.7.14)
and
\[(21.7.17) \quad B_{7,0} = \{e_1 + e_2, e_2 - e_1, e_3 - e_2, e_4 - e_3, e_5 - e_4, e_6 - e_5\}.\]
Thus \(B, B_0\) in the previous section can be expressed as
\[(21.7.18) \quad B = B_7 \cup \{e_7 - e_6\}, \quad B_0 = B_{7,0} \cup \{e_7 - e_6\}.\]
It is easy to see that \(B_{7,0}\) is a basis for \(V_{7,0}\), and in fact \(B_{7,0}\) is a base for \(A_{7,0}\). This is similar to the base discussed in Section 20.12, ordering the standard basis vectors in \(V_{7,0}\) the other way, as before.

Observe that \(B_{7,0}\) is a basis for \(V_{7,0}\), because \(B_{7,0}\) is a basis for \(V_{7,0}\). More precisely, \(B_{7,0}\) is a base for \(A_{7,0}\), as in Section 20.15.

This is how the root system of type \(E_7\) is described on p65 of [13], and p41 of [23]. The maximal root in \(A_{7}\) with respect to \(B_{7}\) is
\[(21.7.19) \quad 2((1/2)(e_1 + e_8 - (e_2 + \cdots + e_7)) + 2(e_1 + e_2) + 3(e_2 - e_1) + 4(e_3 - e_2) + 3(e_4 - e_3) + 2(e_5 - e_4) + (e_6 - e_5)) = e_8 - e_7,\]
as in Table 2 on p66 of [13].

The Coxeter graph of \(A_{7}\) with respect to \(B_{7}\) corresponds to a subgraph of the Coxeter graph of \(A\) with respect to \(B\), and similarly for the Dynkin diagram, as in Section 20.15. These can be depicted as on p58 of [13], or p37, 39 of [23].

If \(T\) is an automorphism of \(A_{7}\) as a root system in \(V_{7}\), and if \(T(B_{7}) = B_{7}\), then \(T\) preserves the Cartan matrix of \(A_{7}\) with respect to \(B_{7}\), as in Section 20.2. This implies that \(T\) preserves the Coxeter graph of \(A\) with respect to \(B\) too. One can verify that this can only happen when \(T\) is the identity mapping on \(V_{7}\), as in Table 1 on p66 of [13], and on p39 of [23].

### 21.8 Type \(E_6\)

We continue with the same notation and hypotheses as in the previous two sections. Remember that \(V = \mathbb{R}^8\), and put
\[(21.8.1) \quad V_6 = \{v \in V : v_6 = v_7 = -v_8\},\]
which is a linear subspace of \(V_7\). Note that \((1/2)(e_1 + e_8 - (e_2 + \cdots + e_7)) \in V_6\). Put
\[(21.8.2) \quad V_{6,0} = V_6 \cap V_0 = \{v \in V : v_6 = v_7 = v_8 = 0\},\]
which is a linear subspace of \(V_{7,0}\).

The elements of
\[(21.8.3) \quad A_1 \cap V_6 = \{\alpha \in \mathbb{Z}^8 \cap V_6 : (\alpha, \alpha) = 2\}\]
are of the form \(\pm e_j \pm e_l\) with \(1 \leq j \neq l \leq 5\). Thus
\[(21.8.4) \quad A_1 \cap V_6 = A_1 \cap V_{6,0},\]
and the linear span of (21.8.4) is $V_{6,0}$. More precisely, (21.8.4) is a reduced root system in $V_{6,0}$, which is isomorphic to the one in Section 20.12 with $n = 5$.

Consider

(21.8.5) $A_6 = A \cap V_6 = \{ \alpha \in \tilde{L}_{8,e} \cap V_6 : (\alpha, \alpha) = 2 \}$.

Note that $A_6 \subseteq A_7$, and that

(21.8.6) $A_6 \cap L_{8,e} = (A \cap L_{8,e}) \cap V_6 = A_1 \cap V_6$.

Of course, $A_6$ is the union of (21.8.6) and

(21.8.7) $A_6 \cap (L_{8,e} + (1/2, -1/2, \ldots, -1/2, 1/2))$

$= (A \cap (L_{8,e} + (1/2, \ldots, 1/2))) \cap V_6$.

An element $\alpha$ of $R^8$ is in (21.8.7) if and only if $\alpha_j = \pm 1/2$ for every $j = 1, \ldots, 8$, $\alpha_j = 1/2$ for an even number of $j$, and $\alpha_6 = \alpha_7 = -\alpha_8$. In particular, $(1/2) (e_1 + e_8 - (e_2 + \cdots + e_7))$ is an element of (21.8.7). It follows that $V_6$ is spanned by $A_6$, because $V_{6,0}$ is spanned by (21.8.6), as in the preceding paragraph. This implies that $A_6$ is a root system in $V_6$, as in Section 20.6. This root system is reduced, and said to be of type $E_6$.

Put

(21.8.8) $A_{6,0} = A_6 \cap V_{6,0} = A \cap V_{6,0}$.

Equivalently,

(21.8.9) $A_{6,0} = A_1 \cap V_{6,0}$,

because $\tilde{L}_{8,e} \cap V_{6,0} = L_{8,e} \cap V_{6,0}$.

Put

(21.8.10) $B_6 = \{(1/2) (e_1 + e_8 - (e_2 + \cdots + e_7)), e_1 + e_2, e_2 - e_1, e_3 - e_2, e_4 - e_3, e_5 - e_4 \}$

and

(21.8.11) $B_{6,0} = \{ e_1 + e_2, e_2 - e_1, e_3 - e_2, e_4 - e_3, e_5 - e_4 \}$.

This means that

(21.8.12) $B_7 = B_6 \cup \{ e_6 - e_5 \}, \quad B_{7,0} = B_{6,0} \cup \{ e_6 - e_5 \}$.

Observe that $B_{6,0}$ is a basis for $V_{6,0}$, and in fact a base for $A_{6,0}$. This is similar to the base discussed in Section 20.12, as before.

It is easy to see that $B_6$ is a basis for $V_6$, because $B_{6,0}$ is a basis for $V_{6,0}$. In fact, $B_6$ is a base for $A_6$, because $B_7$ is a base for $A_7$, as in Section 20.15. Of course, one could also use the fact that $B$ is a base for $A$ here. The root system of type $E_6$ is described in this way on p65 of [13], and p41 of [23].

The maximal root in $A_6$ with respect to $B_6$ is

(21.8.13) $(1/2) (e_1 + e_8 - (e_2 + \cdots + e_7)) + 2 (e_1 + e_2)$

$+ 2 (e_2 - e_1) + 3 (e_3 - e_2) + 2 (e_4 - e_3) + (e_5 - e_4)$

$= (1/2) (e_1 + e_2 + e_3 + e_4 + e_5 - e_6 - e_7 + e_8)$,
as in Table 2 on p66 of [13].

The Coxeter graph of $A_6$ with respect to $B_6$ corresponds to a subgraph of the Coxeter graph of $A_7$ with respect to $B_7$, and similarly for the Dynkin diagram, as in Section 20.15. One can also look at these in terms of their analogues for $A$ and $B$. The Coxeter graph and Dynkin diagram of $A_6$ with respect to $B_6$ may be depicted as on p58 of [13], or p37, 39 of [23].

If $T$ is an automorphism of $A_6$ as a root system in $V_6$, and $T(B_6) = B_6$, then $T$ preserves the Cartan matrix of $A_6$ with respect to $B_6$, and hence the Coxeter graph. There is exactly one automorphism $T$ with this property, other than the identity mapping, which interchanges some elements of $B_6$, and satisfies $T^2 = I$. This is mentioned in Table 1 on p66 of [13], and on p39 of [23].

21.9 Irreducible root systems

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $A$ be an irreducible reduced root system in $V$. It is well known that $A$ is isomorphic to one of the root systems that has been described previously. More precisely, $A$ is isomorphic to one of the following:

- (21.9.1) a root system of type $A_n$, with $n \geq 1$, as in Section 20.10;
- (21.9.2) a root system of type $B_n$, with $n \geq 2$, as in Section 20.11;
- (21.9.3) a root system of type $C_n$, with $n \geq 3$, as in Section 20.11;
- (21.9.4) a root system of type $D_n$, with $n \geq 4$, as in Section 20.12;
- (21.9.5) a root system of type $E_6$, $E_7$, or $E_8$, as in Sections 21.6 – 21.8;
- (21.9.6) a root system of type $F_4$, as in Section 21.4;
- (21.9.7) a root system of type $G_2$, as in Sections 21.2 and 21.3.

Equivalently, if $B$ is a base for $A$, then the corresponding Dynkin diagram is isomorphic to one of the Dynkin diagrams in this list. This is the theorem on p57f of [13], which corresponds to Theorem 4 on p38 of [23]. The restrictions on $n$ in the first four types ensures that there are no repetitions. If $A$ is irreducible but not reduced, then it is well known that $A$ is isomorphic to a root system of type $BC_n$, with $n \geq 1$, as in Section 20.11. This corresponds to Exercise 3 on p66 of [13], and is also mentioned on p41 of [23].

Suppose that $A$ is an irreducible reduced root system again, and let $B$ be a base for $A$. If $T$ is an automorphism of $A$ that maps $B$ onto itself, then $T$ preserves the Cartan matrix of $A$ with respect to $B$, as in Section 20.2. Conversely, any one-to-one mapping from $B$ onto itself that preserves the Cartan matrix corresponds to a unique automorphism $T$ of $A$ that maps $B$ onto itself, as before. Equivalently, automorphisms of $A$ that map $B$ onto itself correspond to automorphisms of the Dynkin diagram of $A$ with respect to $B$.

Remember that there are pairs of vertices in the Coxeter graph of $A$ with respect to $B$ with multiple edges in the cases (21.9.2), (21.9.3), (21.9.6), and (21.9.7). In these cases, the only automorphism of the Dynkin diagram is the
identity mapping, as before. This implies that the automorphism group of $A$ is the same as the Weyl group of $A$.

In the cases (21.9.1), (21.9.4), and (21.9.5), there is at most one edge between any pair of vertices in the Coxeter graph. This means that the Coxeter graph determines the Cartan matrix of $A$ with respect to $B$, or equivalently the Dynkin diagram. In these cases, automorphisms of $A$ that map $B$ onto itself correspond exactly to automorphisms of the Coxeter graph.

If $A$ is of type $A_1$, $E_7$, or $E_8$, then the only automorphism of the Coxeter graph is the identity mapping. Thus the automorphism group of $A$ is the same as the Weyl group of $A$ in these cases.

## 21.10 Connected Coxeter graphs

Let $V$ be a vector space over the real numbers of positive finite dimension, let $A$ be a root system in $V$, and let $B$ be a base for $A$. Also let $(\cdot, \cdot)$ be an inner product on $V$ that is invariant under the Weyl group of $A$, and let $\| \cdot \|$ be the corresponding norm on $V$. If $\alpha \in B$, then put

$$\epsilon_\alpha = \alpha/\|\alpha\|.$$ (21.10.1)

If $\alpha, \beta \in B$ and $\alpha \neq \beta$, then

$$\langle \epsilon_\alpha, \epsilon_\beta \rangle = (\alpha, \beta) \|\alpha\|^{-1} \|\beta\|^{-1} \leq 0.$$ (21.10.2)

where the second step is as in Section 19.11. In this case,

$$4 \langle \epsilon_\alpha, \epsilon_\beta \rangle^2 = 4 (\alpha, \beta)^2 \|\alpha\|^{-2} \|\beta\|^{-2}$$ (21.10.3)

is the same as the number of edges between $\alpha$ and $\beta$ in the Coxeter graph of $A$ with respect to $B$, as in Section 20.3.

Now let $B$ be a nonempty finite set, and let $\Gamma$ be a Coxeter graph, with $B$ as the set of vertices in $\Gamma$. Thus every pair of distinct elements of $B$ is connected by 0, 1, 2, or 3 edges in $\Gamma$. Let $V$ be a vector space over the real numbers, with dimension equal to the number of elements of $B$. Suppose that for each $\alpha \in B$, $\epsilon_\alpha$ is a nonzero element of $V$, and that the collection of $\epsilon_\alpha$, $\alpha \in B$, is a basis for $V$. Under these conditions, there is a unique symmetric bilinear form $(v, w)$ on $V$ with the following properties. First,

$$\langle \epsilon_\alpha, \epsilon_\alpha \rangle = 1$$ (21.10.4)

for every $\alpha \in B$. Second, if $\alpha, \beta \in B$ and $\alpha \neq \beta$, then

$$\langle \epsilon_\alpha, \epsilon_\beta \rangle \leq 0$$ (21.10.5)

and

$$4 \langle \epsilon_\alpha, \epsilon_\beta \rangle^2 = \text{ the number of edges between } \alpha \text{ and } \beta \text{ in } \Gamma.$$ (21.10.6)
We may also be interested in situations where

\[(v, w)\text{ is positive definite on } V.\]

If \(\Gamma\) is the Coxeter graph of a root system, then we can take \(e_\alpha\) to be as in (21.10.1). Thus (21.10.7) holds in this situation, as on p37 of [23]. The properties of the \(e_\alpha\)'s in the preceding paragraph, including (21.10.7), correspond to the definition of an admissible set of vectors in a finite-dimensional inner product space on p60 of [13].

If \(\Gamma\) is a connected Coxeter graph for which (21.10.7) holds, then it is well known that \(\Gamma\) isomorphic to the Coxeter graph of one of the root systems (21.9.1) - (21.9.7). More precisely, the Coxeter graphs associated to root systems of type \(B_n\) and \(C_n\) are the same, and otherwise these Coxeter graphs are distinct. In particular, Coxeter graphs of irreducible root systems are of this form, as in Theorem 3 on p37 of [23]. This is shown in the proof of the theorem on p57f of [13]. Using this, it is not difficult to obtain the analogous statement for Dynkin diagrams, as on p62f of [13], and p39 of [23].

21.11 Lengths in the Weyl group

Let \(V\) be a vector space over the real numbers of positive finite dimension, let \(A\) be a reduced root system in \(V\), and let \(B\) be a base for \(A\). If \(\alpha \in A\), then \(\sigma_\alpha\) denotes the symmetry on \(V\) with vector \(\alpha\) that maps \(A\) onto itself, as usual.

Remember that the Weyl group \(W\) of \(A\) is generated by the symmetries \(\sigma_\beta\) with \(\beta \in B\), as in Section 19.14.

Suppose that \(\sigma \in W\) can be expressed as

\[(21.11.1) \quad \sigma = \sigma_{\beta_1} \circ \cdots \circ \sigma_{\beta_r}\]

for some \(\beta_1, \ldots, \beta_r \in B\), where \(r \geq 0\) is as small as possible. In this case, the expression on the right side of (21.11.1) is said to be reduced, and the length of \(\sigma\) with respect to \(B\) is defined by

\[(21.11.2) \quad \text{length}(\sigma) = r,\]

as on p51 of [13]. More precisely, the right side of (21.11.1) is interpreted as being the identity mapping on \(V\) when \(r = 0\), so that \(\text{length}(\sigma) = 0\) exactly when \(\sigma\) is the identity mapping.

Let \(A^+ = A^{B,+}\) be the set of \(\alpha \in A\) that can be expressed as a linear combination of elements of \(B\) whose coefficients are nonnegative integers, so that \(A = A^+ \cup (-A^+)\). If \(\sigma \in W\), then put

\[(21.11.3) \quad n(\sigma) = \# \{ \alpha \in A^+ : \sigma(\alpha) \in -A^+ \},\]

which is to say the number of elements of the set on the right, as on p51 of [13]. Lemma A at the top of p52 of [13] states that

\[(21.11.4) \quad \text{length}(\sigma) = n(\sigma)\]
for every \( \sigma \in W \). Of course, if \( \text{length}(\sigma) = 0 \), then \( \sigma \) is the identity mapping on \( V \), and \( n(\sigma) = 0 \) too.

Suppose now that \( \text{length}(\sigma) > 0 \), and that the statement holds for elements of \( W \) with smaller length. Let (21.11.1) be a reduced expression for \( \sigma \), so that (21.11.2) holds. Under these conditions, we have that

\[
\sigma(\beta_r) \in -A^+;
\]

as mentioned just after (20.1.8). Equivalently,

\[
(\sigma_{\beta_1} \circ \cdots \circ \sigma_{\beta_{r-1}})(\beta_r) \in A^+,
\]

because \( \sigma_{\beta_r}(\beta_r) = -\beta_r \).

Remember that \( \sigma_{\beta_r} \) maps \( A^+ \) onto itself, as in (19.12.8). Using this, we get that

\[
n(\sigma_{\beta_1} \circ \cdots \circ \sigma_{\beta_{r-1}}) = n(\sigma) - 1.
\]

More precisely, the number of \( \alpha \in A^+ \setminus \{\beta_r\} \) that are mapped into \( -A^+ \) by \( \sigma_{\beta_1} \circ \cdots \circ \sigma_{\beta_{r-1}} \) is the same as for \( \sigma \), because \( \sigma_{\beta_r} \) maps \( A^+ \setminus \{\beta_r\} \) onto itself.

The case where \( \alpha = \beta_r \) corresponds to the remarks in the preceding paragraph.

The length of \( \sigma_{\beta_1} \circ \cdots \circ \sigma_{\beta_{r-1}} \) with respect to \( B \) is clearly less than or equal to \( r - 1 \). In fact,

\[
\text{length}(\sigma_{\beta_1} \circ \cdots \circ \sigma_{\beta_{r-1}}) = r - 1 = \text{length}(\sigma) - 1,
\]

because otherwise \( \sigma \) would have length less than \( r \) with respect to \( B \). Arguing by induction, we have that

\[
\text{length}(\sigma_{\beta_1} \circ \cdots \circ \sigma_{\beta_{r-1}}) = n(\sigma_{\beta_1} \circ \cdots \circ \sigma_{\beta_{r-1}}).
\]

Thus (21.11.4) follows from (21.11.7), (21.11.8), and (21.11.9), as desired.

If \( \sigma \in W \) and \( \beta \in B \), then

\[
n(\sigma \circ \sigma_\beta) = n(\sigma) \pm 1.
\]

Indeed, the number of \( \alpha \in A^+ \setminus \{\beta\} \) that are mapped into \( -A^+ \) by \( \sigma \circ \sigma_\beta \) is the same as for \( \sigma \circ \beta \); because \( \sigma_\beta \) maps \( A \setminus \{\beta\} \) onto itself, as before. Of course,

\[
(\sigma \circ \sigma_\beta)(\beta) = \sigma(\beta_\beta(\beta)) = -\sigma(\beta),
\]

so that exactly one of \( \sigma(\beta) \) and \( (\sigma \circ \sigma_\beta)(\beta) \) is in \( A^+ \). This implies (21.11.10).

Exercise 5 on p54 of [13] states that if \( \sigma \in W \) can be expressed as the composition of \( r \) symmetries associated to elements of \( B \), then \( r \) has the same parity as \( \text{length}(\sigma) \). More precisely, it is easy to see that \( r \) has the same parity as \( n(\sigma) \), by (21.11.11). Thus \( r \) has the same parity as \( \text{length}(\sigma) \), by (21.11.4).

Exercise 6 on p54 of [13] says that

\[
\sigma \mapsto (-1)^{\text{length}(\sigma)}
\]
defines a group homomorphism from \( W \) into the multiplicative group \( \{ \pm 1 \} \). This follows from the fact that if \( \sigma \in W \) can be expressed as the composition of \( r \) symmetries associated to elements of \( \beta \), then
\[
(21.11.13) \quad (-1)^r = (-1)^{\text{length}(\sigma)},
\]
as in the preceding paragraph. Alternatively, the determinant of any symmetry on \( V \) is equal to \(-1\), so that (21.11.12) is the same as the restriction of the determinant to \( W \).

### 21.12 Nonnegative linear functionals

Let us continue with the same notation and hypotheses as in the previous section. Note that
\[
(21.12.1) \quad \text{length}(\sigma^{-1}) = \text{length}(\sigma)
\]
for every \( \sigma \in W \), because the inverse of a symmetry on \( V \) is itself. Similarly,
\[
(21.12.2) \quad n(\sigma^{-1}) = n(\sigma)
\]
for every \( \sigma \in W \). This can be verified directly from the definition (21.11.3) of \( n(\sigma) \), and it is compatible with (21.11.4) and (21.12.1) as well.

Let us say that a linear functional \( \lambda \) on \( V \) is nonnegative with respect to \( B \) if
\[
(21.12.3) \quad \lambda(\beta) \geq 0 \quad \text{for every } \beta \in B.
\]
Of course, this implies that
\[
(21.12.4) \quad \lambda(\alpha) \geq 0 \quad \text{for every } \alpha \in A^+.
\]

Remember that the Weyl chamber associated to \( B \) consists of the \( \lambda \in V' \) such that \( \lambda(\beta) > 0 \) for every \( \beta \in B \), as in Section 19.15. The set of linear functionals on \( V \) that are nonnegative with respect to \( B \) is the same as the closure in \( V' \) of the Weyl chamber associated to \( B \). This uses the topology on \( V' \) that corresponds to the standard topology on a Euclidean space of the same dimension, using any linear isomorphism.

Suppose that \( \lambda \in V' \) is nonnegative with respect to \( B \), \( \sigma \in W \), and \( \lambda \circ \sigma \) is nonnegative with respect to \( B \) too. Under these conditions, \( \sigma \) can be expressed as the composition of finitely many symmetries on \( V \) with vector in \( B \) that map \( A \) onto itself, in such a way that \( \lambda \) is invariant under each of these symmetries. In particular,
\[
(21.12.5) \quad \lambda \circ \sigma = \lambda.
\]
This corresponds to Lemma B on p52 of [13].

If the length of \( \sigma \) with respect to \( B \) is 0, then \( \sigma \) is the identity mapping on \( V \), and the statement in the preceding paragraph is trivial. Suppose that \( \text{length}(\sigma) > 0 \), which is the same as saying that \( \text{length}(\sigma^{-1}) > 0 \). Thus \( n(\sigma^{-1}) \)
is positive, by (21.11.4), so that there is an \( \alpha \in A^+ \) such that \( \sigma^{-1}(\alpha) \in -A^+ \), by the definition (21.11.3) of \( n(\sigma^{-1}) \). More precisely, there is a \( \beta \in B \) such that

\[
\sigma^{-1}(\beta) \in -A^+, \tag{21.12.6}
\]

because otherwise \( \sigma^{-1} \) would map \( A^+ \) into itself. Alternatively, this can be obtained as in (20.1.8).

Because \( \lambda \circ \sigma \) is nonnegative with respect to \( B \), we have that

\[
\lambda(\beta) = (\lambda \circ \sigma)(\sigma^{-1}(\beta)) \leq 0. \tag{21.12.7}
\]

This implies that

\[
\lambda(\beta) = 0, \tag{21.12.8}
\]

because \( \lambda(\beta) \geq 0 \), by hypothesis. It follows that

\[
\lambda \circ \sigma_\beta = \lambda \tag{21.12.9}
\]

on \( V \), because \( \sigma_\beta(v) - v \) is a multiple of \( \beta \) for every \( v \in V \).

Let us check that

\[
\text{length}(\sigma^{-1} \circ \sigma_\beta) = \text{length}(\sigma^{-1}) - 1. \tag{21.12.10}
\]

Equivalently, this means that

\[
n(\sigma^{-1} \circ \sigma_\beta) = n(\sigma^{-1}) - 1, \tag{21.12.11}
\]

by (21.11.4). Remember that \( \sigma_\beta \) maps \( A^+ \setminus \{\beta\} \) onto itself, as in (19.12.8). One can use this to get (21.12.11), because of (21.12.6) and the fact that \( \sigma_\beta(\beta) = -\beta \).

Alternatively, if \( \beta \) is obtained from \( \sigma^{-1} \) as in (20.1.8), then (21.12.10) holds by construction, as in (21.11.8).

Observe that

\[
\lambda \circ (\sigma_\beta \circ \sigma) = (\lambda \circ \sigma_\beta) \circ \sigma = \lambda \circ \sigma, \tag{21.12.12}
\]

by (21.12.9). Thus \( \lambda \circ (\sigma_\beta \circ \sigma) \) is nonnegative with respect to \( B \), by hypothesis. We also have that

\[
\text{length}(\sigma_\beta \circ \sigma) = \text{length}(\sigma) - 1, \tag{21.12.13}
\]

by (21.12.1), (21.12.10), and the fact that \( \sigma_\beta \) is its own inverse. This permits us to use induction to obtain the initial statement about \( \sigma \).

If \( \lambda \) is strictly positive on \( B \), then \( \sigma \) has to be the identity mapping on \( V \). More precisely, (21.12.7) cannot hold in this situation. This corresponds to Exercise 12 on p55 of [13].

Alternatively, if \( \lambda \) is strictly positive on \( B \), then \( B \) is the same as the base \( B_\lambda \) for \( A \) associated to \( \lambda \) as in Section 19.11. If \( \sigma \in W \), then \( \sigma(B) \) is a base for \( A \) too. If \( \lambda \circ \sigma \) is strictly positive on \( B \), then \( \lambda \) is strictly positive on \( \sigma(B) \). This implies that \( \sigma(B) = B_\lambda \), as before, so that \( \sigma(B) = B \). It follows that \( \sigma \) is the identity mapping on \( V \), as in Section 20.1.

Remember that \( \tau \in V' \) is said to be regular with respect to \( A \) if \( \tau(\alpha) \neq 0 \) for each \( \alpha \in A \), as in Section 19.15. In this case, there is a unique base \( B_\tau \) for
A on which $\tau$ is strictly positive, as in Section 19.11. If $\sigma \in W$ and $\tau \circ \sigma = \tau$, then $\tau$ is strictly positive on $\sigma(B_\tau)$, so that $\sigma(B_\tau) = B_\tau$. This implies that $\sigma$ is the identity mapping on $V$, as before.

Let $(\cdot, \cdot)$ be an inner product on $V$ that is invariant under $W$. If $u \in V$, then let $\mu_u(v) = (v, u)$ be the corresponding linear functional on $V$ with respect to $(\cdot, \cdot)$. If $\sigma$ is an orthogonal transformation on $V$ with respect to $(\cdot, \cdot)$, then

$$
\mu_u(\sigma(v)) = (\sigma(v), u) = (v, \sigma^{-1}(u)) = \mu_{\sigma^{-1}(u)}(v)
$$

for every $v \in V$, so that $\mu_u \circ \sigma = \mu_{\sigma^{-1}(u)}$ on $V$. In particular, this holds when $\sigma \in W$. If $\sigma(u) = u$, then $\sigma^{-1}(u) = u$, and $\mu_u \circ \sigma = \mu_u$.

Note that $\mu_u$ is regular with respect to $A$ exactly when $u$ is not orthogonal to any element of $A$ with respect to $(\cdot, \cdot)$. In this case, if $\sigma(u) = u$, then we get that $\sigma$ is the identity mapping on $V$, because $\mu_u \circ \sigma = \sigma$, as before.

If $\sigma \in W$ is a reflection on $V$ with respect to $(\cdot, \cdot)$, then $\sigma$ is the reflection associated to an element of $A$, as in Exercise 13 on p55 of [13]. To see this, it suffices to show that the hyperplane in $V$ of vectors fixed by $\sigma$ is contained in the hyperplane of vectors fixed by $\sigma_\alpha$ for some $\alpha \in A$. Otherwise, for each $\alpha \in A$, the intersection of the hyperplane in $V$ of vectors fixed by $\sigma$ with the hyperplane of vectors fixed by $\sigma_\alpha$ is a hyperplane in the hyperplane of vectors fixed by $\sigma$. Because $A$ has only finitely many elements, there is an element $u$ of the hyperplane of vectors fixed by $\sigma$ that is not in the hyperplane of vectors fixed by $\sigma_\alpha$ for any $\alpha \in A$, as in Section 18.12.

Equivalently, this means that $u$ is not orthogonal to any $\alpha \in A$ with respect to $(\cdot, \cdot)$. By construction, $\sigma(u) = u$, so that $\sigma$ is the identity mapping on $V$, as before. This contradicts the hypothesis that $\sigma$ be a reflection on $V$.

21.13 The mapping $\alpha \mapsto -\alpha$

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $A$ be a root system in $V$. It is easy to see that the mapping

$$
\alpha \mapsto -\alpha
$$

on $V$ is an automorphism of $A$, as in the first part of Exercise 6 on p67 of [13]. The second part of the exercise is to try to decide for which irreducible reduced root systems (21.13.1) is an element of the Weyl group of $A$. Of course, this holds automatically when the automorphism group of $A$ is the same as the Weyl group of $A$. Remember that this happens for root systems of type $B_n$ and $C_n$, as well as types $A_1$, $F_4$, $G_2$, $E_7$, and $E_8$, as in Section 21.9.

If $n \geq 2$, then (21.13.1) is not in the Weyl group of the root system of type $A_n$. This can be seen using the description of the Weyl group in Section 20.10.

The Weyl group of root systems of type $D_\alpha$ consist of linear mappings on $V = \mathbb{R}^n$ that permute the coordinates, and multiply an even number of coordinates by $-1$, as in Section 20.12. Thus (21.13.1) is an element of the Weyl group exactly when $n$ is even.
To deal with type $E_6$, let us take $V = \mathbb{R}^8$, equipped with the standard inner product $(\cdot, \cdot)$, and put

\begin{align}
(21.13.2) & \quad V_6 = \{v \in V : v_6 = v_7 = -v_8\}, \\
(21.13.3) & \quad V_7 = \{v \in V : v_7 = -v_8\},
\end{align}

as in Sections 21.7 and 21.6. Consider the line

\begin{align}
(21.13.4) & \quad Z = \{v \in V : v_1 = \cdots = v_5 = 0, v_7 = -v_6/2 = -v_8\}
\end{align}

in $V_7$. If $v \in V_6$ and $z \in Z$, then

\begin{align}
(21.13.5) & \quad (v, z) = \sum_{j=1}^{n} v_j z_j = v_6 z_6 + v_7 z_7 + v_8 z_8
\end{align}

Thus $Z$ is the orthogonal complement of $V_6$ in $V_7$, with respect to the standard inner product on $V$.

Let $A_6$ and $A_7$ be the root systems in $V_6$ and $V_7$ of types $E_6$ and $E_7$, respectively, as in Sections 21.8 and 21.7. If $\alpha \in A_6$, then the reflection on $V_6$ associated to $\alpha$ is the same as the restriction to $V_6$ of the reflection on $V_7$ associated to $\alpha$. The restriction to $Z$ of the reflection on $V_7$ associated to $\alpha \in V_6$ is the identity mapping on $Z$, because $Z$ is orthogonal to $\alpha$.

Remember that $A_6 = A_7 \cap V_6 \subseteq A_7$, by construction. If $T$ is a one-to-one linear mapping from $V_6$ onto itself, then let $\tilde{T}$ be the extension of $T$ to a one-to-one linear mapping from $V_7$ onto itself whose restriction to $Z$ is the identity mapping on $Z$. If $T$ is in the Weyl group of $A_6$, then it follows that $\tilde{T}$ is in the Weyl group of $A_7$.

Consider the mapping $T$ from $V_6$ into itself defined by $T(v) = -v$ for every $v \in V_6$. If $T$ is in the Weyl group of $A_6$, then $\tilde{T}$ is in the Weyl group of $A_7$, as before. Remember that the mapping on $V_7$ that sends $v \in V_7$ to $-v$ is in the Weyl group of $A_7$, as mentioned earlier. It follows that the composition of this mapping with $T$ is in the Weyl group of $A_7$ as well.

The composed mapping on $V_7$ just mentioned sends every element of $V_6$ to itself, and sends $z \in Z$ to $-z$. Note that this is a reflection on $V_7$. If this composed mapping is an element of the Weyl group of $A_7$, then it is the reflection associated to an element of $A_7$, as in the previous section. This is not possible, because there are no elements of $A_7$ in $Z$. Thus the mapping $v \mapsto -v$ on $V_6$ is not an element of the Weyl group of $A_6$.

21.14 Another mapping on $V$

Let $V$ be a vector space over the real numbers of positive finite dimension again, and let $A$ be a reduced root system in $V$. Also let $B$ be a base for $A$, and observe
that $-B$ is a base for $A$ as well. It follows that there is an element $\sigma$ of the Weyl group $W$ of $A$ such that

$$\sigma(B) = -B,$$

as in Section 19.14. In fact, $\sigma$ is unique, as in Section 20.1.

Let $A^+ = A^{B,+}$ be the set of $\alpha \in A$ that can be expressed as linear combinations of elements of $B$ whose coefficients are nonnegative integers, as before. Of course, (21.14.1) implies that

$$\sigma(A^+) = -A^+. \quad (21.14.2)$$

If $n(\sigma)$ is as in (21.11.3), then (21.14.2) is the same as saying that

$$n(\sigma) = \#A^+. \quad (21.14.3)$$

Note that this is the largest possible value of $n(\sigma)$. Remember that this is equal to the length of $\sigma$ with respect to $B$, as in (21.11.4).

If $\sigma$ is any element of $W$ that satisfies (21.14.2), then $\sigma$ satisfies (21.14.1). Indeed, (21.14.2) implies that

$$\sigma(B) \subseteq -A^+ = A^{-B,+}. \quad (21.14.4)$$

This implies that (21.14.1) holds, because $\sigma(B)$ is a base for $A$, as in (19.15.7).

Let $\beta_0 \in B$ be given, and let $V_0$ be the hyperplane in $V$ spanned by $\beta \in B$ with $\beta \neq \beta_0$. If $v \in V$, then $v$ can be expressed in a unique way as a linear combination of elements of $B$, and we let $\tau_0(v) \in \mathbb{R}$ be the coefficient of $\beta_0$ in this expression for $v$. Thus $V_0$ is the same as the kernel of $\tau_0$. If $\beta \in B \setminus \{\beta_0\}$ and $v \in V$, then $\sigma_\beta(v) - v \in V_0$, so that

$$\tau_0 \circ \sigma_\beta = \tau_0 \quad (21.14.5)$$

on $V$. If $\sigma$ is an element of the subgroup of $W$ generated by $\sigma_\beta$ with $\beta \in B \setminus \{\beta_0\}$, then it follows that

$$\tau_0 \circ \sigma = \tau_0 \quad (21.14.6)$$

on $V$ too. In particular, this means that

$$\tau_0(\sigma(\beta_0)) = \tau_0(\beta_0) = 1. \quad (21.14.7)$$

This implies that $\sigma$ does not satisfy (21.14.2).

This corresponds to Exercise 9 on p54 of [13].

## 21.15 Invertibility of Cartan matrices

Let $k$ be a field, and let $V$ be a vector space over $k$ of positive finite dimension $r$. Also let $k^r$ be the space of $r$-tuples of elements of $k$, as usual, which is an $r$-dimensional vector space over $k$ with respect to coordinatewise addition and
21.15. INVERTIBILITY OF CARTAN MATRICES

scalar multiplication. If \( v_1, \ldots, v_r \) is a basis for \( V \), then we get a one-to-one linear mapping from \( k^r \) onto \( V \), which sends \( t = (t_1, \ldots, t_r) \in k^r \) to \( \sum_{j=1}^{r} t_j v_j \). Similarly, if \( \lambda_1, \ldots, \lambda_r \) are linear functionals on \( V \) that form a basis for the dual space \( V' \), then we get a one-to-one mapping from \( V \) onto \( k^r \), which sends \( v \in V \) to \( (\lambda_1(v), \ldots, \lambda_r(v)) \in k^r \). Thus the composition of these two mappings is a one-to-one linear mapping from \( k^r \) onto itself. The composition of these two mappings corresponds to the \( r \times r \) matrix with entries \( \lambda_l(v_j) \), \( j, l = 1, \ldots, r \), with respect to the standard basis in \( k^r \). Thus this matrix is invertible, as an \( r \times r \) matrix with entries in \( k \), under these conditions.

Now let \( V \) be a vector space over the real numbers of positive finite dimension, and let \( A \) be a root system in \( V \). As before, let \( \sigma_\alpha \) be the symmetry on \( V \) with vector \( \alpha \in A \) that maps \( A \) onto itself, and let \( \lambda_\alpha \) be the corresponding linear functional on \( V \). Remember that \( n(\alpha, \beta) = \lambda_\beta(\alpha) \) is an integer for every \( \alpha, \beta \in A \), by the definition of a root system. Let \( \langle \cdot, \cdot \rangle \) be an inner product on \( V \) that is invariant under the Weyl group of \( A \), so that \( \sigma_\alpha \) is the reflection on \( V \) with respect to \( \langle \cdot, \cdot \rangle \) associated to \( \alpha \in A \), and \( \lambda_\alpha(v) = 2 \langle v, \alpha \rangle (\alpha, \alpha)^{-1} \) for every \( v \in V \). Thus \( n(\alpha, \beta) = 2 \langle \alpha, \beta \rangle (\beta, \beta)^{-1} \) for every \( \alpha, \beta \in A \).

Let \( B \) be a base for \( A \), so that \( B \) is a basis for \( V \) in particular. Using the inner product, we can identify \( V \) with its dual space, and it follows that \( \lambda_\beta, \beta \in B \), is a basis for \( V' \). This implies that the Cartan matrix \( n(\alpha, \beta), \alpha, \beta \in B \), is invertible, as before, and as on p55 of [13]. Equivalently, the determinant of the Cartan matrix is nonzero. Of course, the determinant is an integer.

More precisely, the Cartan matrix may be considered as the product of the symmetric matrix \( (\alpha, \beta) \) and the diagonal matrix with diagonal entries \( 2 \langle \beta, \beta \rangle^{-1} \). The symmetric matrix \( (\alpha, \beta) \) is positive definite, because of the positivity of the inner product, and the fact that \( B \) is a basis for \( V \). Thus its determinant is a positive real number. The determinant of the diagonal matrix is positive too, because the diagonal entries of the matrix are positive. This means that the determinant of the Cartan matrix is positive as well.
Chapter 22

Roots and Lie algebras

22.1 A basic situation

Let \( k \) be a field of characteristic 0, and let \((A, [\cdot, \cdot]_A)\) be a Lie algebra over \( k \) with positive finite dimension as a vector space over \( k \). Also let \( B \) be a Lie subalgebra of \( A \) such that every element of \( B \) is ad-diagonalizable as an element of \( A \). This implies that \( B \) is commutative as a Lie algebra, as in Section 17.1.

Let \( B' \) be the dual space of linear functionals on \( B \), as a vector space over \( k \), as usual. If \( \beta \in B' \), then put

\[
A_\beta = \{ x \in A : \text{ad}_w(x) = [w, x]_A = \alpha(w) x \text{ for every } w \in B \},
\]

as before, which is a linear subspace of \( A \). This is the same as the centralizer \( C_A(B) \) of \( B \) in \( A \) when \( \alpha = 0 \). In particular, \( B \subseteq A_0 \), because \( B \) is commutative as a Lie algebra. Let us suppose for the rest of the section that

\[
A_0 = C_A(B) = B
\]

and

\[
Z(A) = \{0\},
\]

where \( Z(A) \) is the center of \( A \) as a Lie algebra.

It follows that \( B \neq \{0\} \), because \( A \neq \{0\} \), by hypothesis. Put

\[
\Phi_B = \{ \alpha \in B' : \alpha \neq 0 \text{ and } A_\alpha \neq \{0\} \},
\]

so that \( \Phi_B \cup \{0\} \) is the same as the set of \( \alpha \in B' \) such that \( A_\alpha \neq \{0\} \), as before. Remember that \( A \) corresponds to the direct sum of the subspaces \( A_\alpha \) with \( \alpha \in \Phi_B \cup \{0\} \), as a vector space over \( k \), and in particular that \( \Phi_B \) has only finitely many elements, as in Section 17.2. If \( x \in A_\alpha \) and \( y \in A_\beta \) for some \( \alpha, \beta \in B' \), then

\[
[x, y]_A \in A_{\alpha+\beta},
\]

as before. We also have that the linear span of \( \Phi_B \) in \( B' \) is equal to \( B' \), because of (22.1.3), as in Section 17.5. Let \( b(\cdot, \cdot) \) be a nondegenerate bilinear form on \( A \),
and suppose that \( b(\cdot, \cdot) \) is associative on \( A \), or equivalently invariant under the adjoint representation on \( A \), as in Sections 6.10 and 7.7. This implies that the restriction of \( b(\cdot, \cdot) \) to (22.1.2) is nondegenerate, as in Section 17.3.

Of course, there is a natural embedding of \( \mathbb{Q} \) into \( k \), because \( k \) has characteristic 0, so that \( B' \) may also be considered as a vector space over \( \mathbb{Q} \). Let \( E_Q \) be the linear subspace of \( B' \), as a vector space over \( \mathbb{Q} \), spanned by \( \Phi_B \), as in Section 17.12. If \( \alpha_1, \ldots, \alpha_n \) are elements of \( \Phi_B \) that form a basis for \( B' \), as a vector space over \( k \), then \( \alpha_1, \ldots, \alpha_n \) form a basis for \( E_Q \), as a vector space over \( \mathbb{Q} \), as in Section 17.10. Using this, we can get a vector space \( E_R \) over \( \mathbb{R} \) of the same dimension \( n \) that contains \( E_Q \), as in Section 17.13. In particular, \( \Phi_B \) is contained in \( E_R \), and the linear span of \( \Phi_B \) in \( E_R \), as a vector space over \( \mathbb{R} \), is equal to \( E_R \).

If \( \alpha \in B' \), then there is a unique \( t_{b, \alpha} \in B \) such that \( \alpha(w) = b(w, t_{b, w}) \) for every \( w \in B \), because \( b(\cdot, \cdot) \) is nondegenerate on \( B \). Remember that \( \alpha(t_{b, \alpha}) = b(t_{b, \alpha}, t_{b, \alpha}) \neq 0 \) when \( \alpha \in \Phi_B \), as in Section 17.6. In this case, we put

\[
(22.1.6) \quad h_\alpha = 2 \alpha(t_{b, \alpha})^{-1} t_{b, \alpha},
\]

as before, so that \( \alpha(h_\alpha) = 2 \). If \( \beta \in \Phi_B \), then \( \beta - \beta(h_\alpha) \alpha \in \Phi_B \), as in Section 17.8. We have also seen that \( \beta(h_\alpha) \) corresponds to an integer in this situation, with respect to the standard embedding of \( \mathbb{Q} \) into \( k \).

Of course,

\[
(22.1.7) \quad \beta \mapsto \beta(h_\alpha)
\]

defines a linear functional on \( B' \), as a vector space over \( k \). If \( \beta \in E_Q \), then \( \beta(h_\alpha) \) is in the image of \( \mathbb{Q} \) in \( k \), by the remarks in the previous paragraphs. Let \( \lambda_\alpha \) be the corresponding linear functional on \( E_Q \), as a vector space over \( \mathbb{Q} \).

This extends to a linear functional on \( E_R \), as a vector space over \( \mathbb{R} \), in a natural way, that we shall also denote by \( \lambda_\alpha \). As before, we have that

\[
(22.1.8) \quad \beta - \lambda_\alpha(\beta) \alpha \in \Phi_B
\]

and \( \lambda_\alpha(\beta) \in \mathbb{Z} \) for every \( \beta \in \Phi_B \).

It follows that \( \Phi_B \) is a root system in \( E_R \), as in Section 19.3. More precisely, if \( \alpha \in \Phi_B \), then the corresponding symmetry on \( E_R \) is defined by

\[
(22.1.9) \quad \sigma_\alpha(\beta) = \beta - \lambda_\alpha(\beta) \alpha
\]

for every \( \beta \in E_R \). Note that the definition of a root system is extended to finite-dimensional vector spaces over the complex numbers in Definition 7 on p41 of [23]. If \( k = \mathbb{C} \), then one can consider \( \Phi_B \) as a root system in \( B' \) as a vector space over \( \mathbb{C} \).

This root system is reduced, in the sense that for each \( \alpha \in \Phi_B \), the only multiples of \( \alpha \) in \( \Phi_B \) are \( \pm \alpha \). This was discussed in Sections 17.7 and 17.8, for multiples of \( \alpha \) by elements of \( k \), and by elements of \( \mathbb{Q} \) in particular. One should also consider multiples of \( \alpha \) by real numbers in \( E_R \), but it is easy to reduce to considering multiples by elements of \( \mathbb{Q} \), because \( \Phi_B \) is contained in \( E_Q \). Of course, in any root system, one only has to consider multiples by \( \pm 2 \) or \( \pm 1/2 \).
Let $(\cdot, \cdot)$ be any inner product on $E_R$ that is invariant under the Weyl group of $\Phi_B$. If $\alpha \in \Phi_B$, then it follows that (22.1.9) is the reflection on $E_R$ associated to $\alpha$ with respect to $(\cdot, \cdot)$. This means that
\[
(22.1.10) \quad \lambda_\alpha(\beta) = 2(\beta, \alpha)(\alpha, \alpha)^{-1}
\]
for every $\beta \in E_R$. In particular, this corresponds to $\beta(h_\alpha)$ when $\beta \in \Phi_B$ too.

Suppose for the moment that $A$ is semisimple as a Lie algebra over $k$, so that we can take $b(\cdot, \cdot)$ to be the Killing form on $A$, as in Section 17.11. This leads to an inner product $(\cdot, \cdot)_{E_Q}$ on $E_Q$, as in Section 17.13. More precisely, a bilinear form $(\cdot, \cdot)_{E_Q}$ was defined on $E_Q$ in Section 17.12, which was extended to $E_R$. If $\alpha, \beta \in \Phi_B$, then
\[
(22.1.11) \quad \lambda_\alpha(\beta) = 2(\alpha, \beta)_{E_Q}(\alpha, \alpha)^{-1}_{E_Q},
\]
by (17.10.8) and the definition of $(\cdot, \cdot)_{E_Q}$. Of course, this implies that
\[
(22.1.12) \quad \lambda_\alpha(\beta) = 2(\alpha, \beta)_{E_R}(\alpha, \alpha)^{-1}_{E_R}.
\]
In fact, this holds for every $\alpha \in \Phi_B$ and $\beta \in E_R$, because $E_R$ is spanned by $\Phi_B$, and both sides are linear in $\beta$. This means that (22.1.9) is the same as the reflection on $E_R$ associated to $\alpha \in \Phi_B$ with respect to $(\cdot, \cdot)_{E_R}$ in this situation. This corresponds to the theorem on p40 of [13].

22.2 Inverse roots and $B$

Let us continue with the same notation and hypotheses as in the previous section, except for the last paragraph. Let $\alpha_1, \ldots, \alpha_n$ be elements of $\Phi_B$ that form a basis for $B'$ as a vector space over $k$, so that $\alpha_1, \ldots, \alpha_n$ form a basis for $E_Q$ as a vector space over $Q$, as before. Thus $\alpha_1, \ldots, \alpha_n$ is a basis for $E_R$ too, as a vector space over $R$, by construction.

Let $E'_Q$ and $E'_R$ be the dual spaces of linear functionals on $E_Q$ and $E_R$, as vector spaces over $Q$ and $R$, respectively. Of course, any linear functional on $E_Q$ or $E_R$ is uniquely determined by its values on $\alpha_1, \ldots, \alpha_n$. Linear functionals on $E_Q$ may take arbitrary values in $Q$ on $\alpha_1, \ldots, \alpha_n$, and linear functionals on $E_R$ may take arbitrary values in $R$ on $\alpha_1, \ldots, \alpha_n$. Every linear functional on $E_Q$ has a natural unique extension to a linear functional on $E_R$. If $\mu$ is a linear functional on $E_R$, and $\mu$ maps $E_Q \subseteq E_R$ into $Q$, then the restriction of $\mu$ to $E_Q$ is a linear functional on $E_Q$. Thus we may identify $E'_Q$ with
\[
(22.2.1) \quad \{\mu \in E'_R : \mu(E_Q) \subseteq Q\}.
\]
This is a linear subspace of $E'_R$, as a vector space over $Q$.

If $w \in B$, then put
\[
(22.2.2) \quad T(w) = (\alpha_1(w), \ldots, \alpha_n(w)),
\]
which is an element of the space $k^n$ of $n$-tuples of elements of $k$. This defines a linear mapping from $B$ into $k^n$, as a vector space over $k$ with respect to coordinatewise addition and scalar multiplication, because $\alpha_j$ is a linear functional on $B$ as a vector space over $k$ for each $j = 1, \ldots, n$. More precisely, $T$ is a one-to-one mapping from $B$ onto $k^n$, because $\alpha_1, \ldots, \alpha_n$ is a basis for $B'$, as a vector space over $k$.

Remember that $k$ has characteristic 0, so that there is a standard embedding of $\mathbb{Q}$ into $k$. Let us identify $\mathbb{Q}$ with its image in $k$ under the standard embedding, so that $\mathbb{Q}^n$ may be identified with a subset of $k^n$. Of course, $\mathbb{Q}^n$ is a linear subspace of $k^n$, as a vector space over $\mathbb{Q}$. Put

\begin{equation}
C_\mathbb{Q} = T^{-1}(\mathbb{Q}^n),
\end{equation}

which is a linear subspace of $B$, as a vector space over $\mathbb{Q}$. Equivalently, $C_\mathbb{Q}$ consists of the $w \in B$ such that $\alpha_j(w) \in \mathbb{Q}$ for every $j = 1, \ldots, n$. This is the same as

\begin{equation}
C_\mathbb{Q} = \{ w \in B : \beta(w) \in \mathbb{Q} \text{ for every } \beta \in E_\mathbb{Q} \},
\end{equation}

by the definition of $E_\mathbb{Q}$. Note that the dimension of $C_\mathbb{Q}$, as a vector space over $\mathbb{Q}$, is equal to $n$, and that the linear span of $C_\mathbb{Q}$ in $B$, as a vector space over $k$, is equal to $B$. The standard basis for $\mathbb{Q}^n$ corresponds to a basis for $C_\mathbb{Q}$, as a vector space over $\mathbb{Q}$, which is also a basis for $B$, as a vector space over $k$.

If $\beta \in E_\mathbb{Q} \subseteq B'$, then $\beta$ is a linear functional on $B$, as a vector space over $k$, that maps $C_\mathbb{Q}$ into $\mathbb{Q}$. This implies that the restriction of $\beta$ to $C_\mathbb{Q}$ is a linear functional on $C_\mathbb{Q}$, as a vector space over $\mathbb{Q}$. Of course, $\beta$ is uniquely determined, as an element of $B'$, by its restriction to $C_\mathbb{Q}$, because the linear span of $C_\mathbb{Q}$ in $B$, as a vector space over $k$, is equal to $B$. It is easy to see that every linear functional on $C_\mathbb{Q}$, as a vector space over $\mathbb{Q}$, corresponds to some $\beta \in E_\mathbb{Q}$ in this way, using $T$. Thus the dual space $C_\mathbb{Q}'$ of linear functionals on $C_\mathbb{Q}$, as a vector space over $\mathbb{Q}$, can be identified with $E_\mathbb{Q}$.

If $w \in B$, then

\begin{equation}
\beta \mapsto \beta(w)
\end{equation}

defines a linear functional on the dual $B'$ of $B$, as a vector space over $k$. This defines a one-to-one linear mapping from $B$ onto the dual of $B'$, as vector spaces over $k$. Similarly, if $w \in C_\mathbb{Q}$, then (22.2.5) defines a linear functional on $E_\mathbb{Q}$, as a vector space over $\mathbb{Q}$. This defines a one-to-one linear mapping from $C_\mathbb{Q}$ onto $E_\mathbb{Q}'$, as vector spaces over $\mathbb{Q}$.

Let $\alpha \in \Phi_B$ be given, and let $h_\alpha \in B$ be as in the previous section. Remember that $\beta(h_\alpha) \in \mathbb{Z}$ for every $\beta \in \Phi_B$, so that

\begin{equation}
h_\alpha \in C_\mathbb{Q}
\end{equation}

in particular. The linear functional $\lambda_\alpha$ on $E_\mathbb{Q}$ that corresponds to $h_\alpha$ as in the preceding paragraph is the same as in the previous section. Note that $\lambda_\alpha$, considered as a linear functional on $E_\mathbb{R}$, is the inverse root associated to $\alpha$, as in Section 19.8.
CHAPTER 22. ROOTS AND LIE ALGEBRAS

Put
\[ \Phi_B' = \{ \lambda_\alpha : \alpha \in \Phi_B \} , \]
which is a subset of \( E'_Q \subseteq E'_R \). As a subset of \( E'_R \), this is the inverse system of \( \Phi_B \), as in Section 19.8.

In [23], \( k = \mathbb{C} \), and the notion of root systems is extended to complex vector spaces, so that \( h_\alpha \) is identified directly with the inverse root of \( \alpha \), as in part (b) of Theorem 2 on p43f.

In [13], the Killing form on a semisimple Lie algebra \( A \) is used to get an inner product on \( E_R \), as in the previous section. Using this, the inverse root of \( \alpha \) corresponds to an element of \( E_Q \), and thus an element of \( B' \). This corresponds to \( h_\alpha \in B \) using the Killing form, as on p43 of [13].

22.3 Semigroups and submodules

Let \( \Xi \) be a commutative semigroup, with the semigroup operation expressed additively. Also let \( \Psi \) be a nonempty subset of \( \Xi \). If \( \alpha \in \Xi \) and \( n \) is a positive integer, then \( n \cdot \alpha \) is the sum of \( n \) \( \alpha \)'s in \( \Xi \), as usual.

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( (A, [\cdot , \cdot]_A) \) be a Lie algebra over \( k \). Suppose that for each \( \alpha \in \Psi \), \( A_\alpha \) is a submodule of \( A \), as a module over \( k \). We ask that \( A \) correspond to the direct sum of the \( A_\alpha \)'s, \( \alpha \in \Psi \), as a module over \( k \), so that every element of \( A \) can be expressed in a unique way as a sum of elements of the \( A_\alpha \)'s. It is convenient to put
\[ A_\alpha = \{0\} \quad \text{when} \quad \alpha \in \Xi \setminus \Psi , \]
so that \( A_\alpha \) is defined as a submodule of \( A \), as a module over \( k \), for every \( \alpha \in \Xi \).

If \( x \in A_\alpha \) and \( y \in A_\beta \) for some \( \alpha, \beta \in \Xi \), then we ask that
\[ [x,y]_A \in A_{\alpha + \beta} . \]
Of course, this holds trivially when \( \alpha \) or \( \beta \) is in \( \Xi \setminus \Psi \), by (22.3.1). Observe that
\[ [x,y]_A = 0 \quad \text{when} \quad \alpha + \beta \notin \Psi , \]
by (22.3.1) and (22.3.2). If \( \Xi \) has an additive identity element 0, then (22.3.2) implies that \( A_0 \) is a Lie subalgebra of \( A \).

If \( x \in A_\alpha \) for some \( \alpha \in \Psi \), then
\[ \text{ad}_x(A_\gamma) \subseteq A_{\alpha + \gamma} \]
for every \( \gamma \in \Xi \). If \( n \in \mathbb{Z}_+ \), then we get that
\[ (\text{ad}_x)^n(A_\gamma) \subseteq A_{n \cdot \alpha + \gamma} , \]
where \( (\text{ad}_x)^n \) is the \( n \)th power of \( \text{ad}_x \) on \( A \), with respect to composition of mappings. Suppose that \( \gamma \in \Psi \) and \( n(\gamma) \in \mathbb{Z}_+ \) have the property that
\[ n(\gamma) \cdot \alpha + \gamma \notin \Psi . \]
This implies that \((\text{ad}_x)^n(\gamma) = 0\) on \(A_\gamma\), and hence that \((\text{ad}_x)^n = 0\) on \(A_\gamma\) for every integer \(n \geq n(\gamma)\). If \(\Psi\) has only finitely many elements, and for each \(\gamma \in \Psi\) there is an \(n(\gamma) \in \mathbb{Z}_+\) such that (22.3.6) holds, then it follows that \((\text{ad}_x)^n = 0\) on \(A\) when \(n \geq \max_{\gamma \in \Psi} n(\gamma)\), so that \(x\) is ad-nilpotent as an element of \(A\).

In some situations, we may have that \((n \cdot \alpha + \gamma)\) is injective as a mapping from \(\mathbb{Z}_+\) into \(\Xi\). In particular, this holds for every \(\gamma \in \Xi\) when \(\Xi\) is a commutative group without torsion, and \(\alpha \neq 0\). Suppose that \(\Psi\) has only finitely many elements, so that the injectivity of (22.3.7) implies that

\[
22.3.7 \quad n \mapsto n \cdot \alpha + \gamma
\]

is injective as a mapping from \(\mathbb{Z}_+\) into \(\Xi\). In particular, this holds for every \(\gamma \in \Xi\) when \(\Xi\) is a commutative group without torsion, and \(\alpha \neq 0\). Suppose that \(\Psi\) has only finitely many elements, so that the injectivity of (22.3.7) implies that

\[
22.3.8 \quad n \cdot \alpha + \gamma \in \Psi
\]

for only finitely many \(n \in \mathbb{Z}_+\). This means that \((\text{ad}_x)^n = 0\) on \(A_\gamma\) for all but finitely many \(n \in \mathbb{Z}_+\), as before. If this holds for every \(\gamma \in \Psi\), then it follows that \((\text{ad}_x)^n = 0\) on \(A\) for all but finitely many \(n \in \mathbb{Z}_+\).

As usual, one may say that \(\Xi\) is a semigroup with cancellation if for every \(\alpha, \beta, \gamma \in \Xi\) with

\[
(22.3.9) \quad \alpha + \gamma = \beta + \gamma,
\]

we have that \(\alpha = \beta\). Note that this holds when \(\Xi\) is a subsemigroup of a group.

Suppose now that \(\Xi\) has an additive identity element 0. If \(\alpha \in \Xi\) satisfies \(n \cdot \alpha = 0\) for some \(n \in \mathbb{Z}_+\), then \(\alpha\) is said to be a torsion element of \(\Xi\). If 0 is the only torsion element of \(\Xi\), then one may say that \(\Xi\) is torsion-free, or that \(\Xi\) has no nontrivial torsion.

Suppose that \(\Xi\) is a semigroup with cancellation. If \(\alpha, \gamma \in \Xi\) and \(\alpha\) is not a torsion element, then it is easy to see that (22.3.8) is injective as a mapping from \(\mathbb{Z}_+\) into \(\Xi\). If \(x \in A_\alpha\) and \(\Psi\) has only finitely many elements, then it follows that \(x\) is ad-nilpotent as an element of \(A\), as before.

### 22.4 Semigroups and algebras

Let \(\Xi\) be a commutative semigroup, where the semigroup operation is expressed additively, and let \(\Psi\) be a nonempty subset of \(\Xi\). Also let \(k\) be a commutative ring with a multiplicative identity element, and let \(A\) be an algebra over \(k\) in the strict sense, where multiplication of \(x, y \in A\) is expressed as \(xy\). Suppose that for each \(\alpha \in \Psi\), \(A_\alpha\) is a submodule of \(A\), as a module over \(k\), and that \(A\) corresponds to the direct sum of the \(A_\alpha\)’s, \(\alpha \in \Psi\), as a module over \(k\). As before, it is convenient to put \(A_\alpha = \{0\}\) when \(\alpha \in \Xi \setminus \Psi\).

Suppose that

\[
(22.4.1) \quad xy \in A_{\alpha+\beta}
\]

when \(x \in A_\alpha\) and \(y \in A_\beta\) for some \(\alpha, \beta \in \Xi\). Thus

\[
(22.4.2) \quad xy = 0
\]
when $\alpha$, $\beta$, or $\alpha + \beta$ is not in $\Psi$. If $\Xi$ has an additive identity element 0, then $A_0$ is a subalgebra of $A$, by (22.4.1). Of course, the remarks in the previous section correspond to the case where $A$ is a Lie algebra over $k$.

Note that $\langle xy, yx \rangle \in A_{\alpha+\beta}$ when $x \in A_\alpha$ and $y \in A_\beta$ for some $\alpha, \beta \in \Xi$, by (22.4.1). If $A$ is an associative algebra over $k$, then $A$ may be considered as a Lie algebra with respect to the corresponding commutator bracket, as usual.

Suppose that $A$ is an associative algebra over $k$, and that $x \in A_\alpha$ for some $\alpha \in \Psi$. If $n \in \mathbb{Z}_+$, then $x^n \in A_{n-\alpha}$.

by (22.4.1). In particular, if $n \cdot \alpha \not\in \Psi$, then $x^n = 0$. If $n \mapsto n \cdot \alpha$ is injective as a mapping from $\mathbb{Z}_+$ into $\Xi$, and $\Psi$ has only finitely many elements, then $n \cdot \alpha \in \Psi$ for only finitely many $n \in \mathbb{Z}_+$, so that $x$ is nilpotent in $A$. If $\Xi$ is a semigroup with cancellations and an additive identity element 0, and if $\alpha$ is not a torsion element of $\Xi$, then $n \mapsto n \cdot \alpha$ is an injective mapping from $\mathbb{Z}_+$ into $\Xi$, as before.

Let $A$ be an algebra over $k$ in the strict sense again. If $0$ is a subset of $\Psi$, then let $A(0)$ be the subset of $A$ consisting of elements that can be expressed as a sum of elements of $A_\alpha$, for finitely many $\alpha \in \Psi$. This is a submodule of $A$, as a module over $k$, which may be interpreted as being $\{0\}$ when $\Psi = \emptyset$, and which corresponds to the direct sum of the $A_\alpha$’s, $\alpha \in \Psi_0$, as a module over $k$. Of course, $A(\Psi_0) = A(\Psi_0 \cap \Psi)$, and one may wish to restrict one’s attention to $\Psi_0 \subseteq \Psi$. However, it is sometimes convenient for $A(\Psi_0)$ to be defined for any $\Psi_0 \subseteq \Xi$.

Suppose that $\Psi_0 \subseteq \Xi$ has the property that

$$\text{if } \alpha, \beta \in \Psi_0 \text{ and } \alpha + \beta \in \Psi, \text{ then } \alpha + \beta \in \Psi_0.$$  

In particular, if $\Xi_0$ is a subsemigroup of $\Xi$, then $\Xi_0 \cap \Psi$ has this property. Under these conditions, $A(\Psi_0)$ is a subalgebra of $A$, as an algebra over $k$ in the strict sense.

Suppose that $\Psi_1 \subseteq \Xi$ has the property that

$$\text{if } \alpha \in \Psi_1 \text{ and } \beta, \alpha + \beta \in \Psi, \text{ then } \alpha + \beta \in \Psi_1.$$  

In this case, $A(\Psi_1)$ is a two-sided ideal in $A$. If $\Xi_1 \subseteq \Xi$ has the property that

$$\text{if } \alpha \in \Xi_1 \text{ and } \beta \in \Xi, \text{ then } \alpha + \beta \in \Xi_1,$$

then $\Xi_1 \cap \Psi$ satisfies (22.4.7).

Similarly, suppose that $\Psi_2 \subseteq \Xi$ satisfies (22.4.6), and that $\Psi_1 \subseteq \Psi_2$ satisfies

$$\text{if } \alpha \in \Psi_1 \text{ and } \beta, \alpha + \beta \in \Psi_2, \text{ then } \alpha + \beta \in \Psi_2.$$  

In this situation, $A(\Psi_1)$ is a two-sided ideal in $A(\Psi_2)$, as an algebra over $k$ in the strict sense.
22.5 Some nilpotency properties

Let us continue with the same notation and hypotheses as in the previous section. In particular, $\Xi$ is a commutative semigroup, and $A$ is an algebra over $k$ in the strict sense. If $\Psi_1, \ldots, \Psi_r \subseteq \Xi$ for some $r \in \mathbb{Z}_+$, then put

$$\Psi_1 + \cdots + \Psi_r = \left\{ \sum_{j=1}^r \alpha_j : \alpha_j \in \Psi_j \text{ for each } j = 1, \ldots, r \right\},$$

as usual. This subset of $\Xi$ may be denoted $\sum_{j=1}^r \Psi_j$ as well. Note that

$$A(\Psi_1) \cdot A(\Psi_2) \subseteq A(\Psi_1 + \Psi_2)$$

for every $\Psi_1, \Psi_2 \subseteq \Xi$, by (22.4.1), where the left side is the set of finite sums of products of elements of $A(\Psi_1)$ and $A(\Psi_2)$, as in Section 9.2.

If $\Psi_0 \subseteq \Xi$, then (22.4.6) may be expressed as

$$(\Psi_0 + \Psi_0) \cap \Psi \subseteq \Psi_0.$$  \hspace{1cm} (22.5.3)

Of course, $\Xi_0 \subseteq \Xi$ is a subsemigroup of $\Xi$ when

$$\Xi_0 + \Xi_0 \subseteq \Xi_0.$$  \hspace{1cm} (22.5.4)

Similarly, if $\Psi_1 \subseteq \Xi$, then (22.4.7) may be expressed as

$$(\Psi_1 + \Psi) \cap \Psi \subseteq \Psi_1.$$  \hspace{1cm} (22.5.5)

If $\Xi_1 \subseteq \Xi$, then (22.4.8) may be expressed as

$$\Xi_1 + \Xi \subseteq \Xi_1.$$  \hspace{1cm} (22.5.6)

If $\Psi_1 \subseteq \Psi_2 \subseteq \Xi$, then (22.4.9) may be expressed as

$$(\Psi_1 + \Psi_2) \cap \Psi_2 \subseteq \Psi_1.$$  \hspace{1cm} (22.5.7)

If $r \in \mathbb{Z}_+$, and $\Psi_j \subseteq \Xi, x_j \in A(\Psi_j)$ for each $j = 1, \ldots, r$, then the product of $x_1, \ldots, x_r$ in $A$, with any ordering or grouping, is an element of $A\left(\sum_{j=1}^r \Psi_j\right)$. In particular, if

$$\left(\sum_{j=1}^r \Psi_j\right) \cap \Psi = \emptyset,$$

then any such product of $x_1, \ldots, x_n$ in $A$ is equal to 0.

Let $\Psi_0$ be a subset of $\Xi$, and put

$$\Psi_0(r) = \sum_{j=1}^r \Psi_0$$

as usual. This subset of $\Xi$ may be denoted $\sum_{j=1}^r \Psi_j$ as well. Note that
for each $r \in \mathbb{Z}_+$, which is the subset of $\Xi$ consisting of sums of $r$ elements of $\Psi_0$. If $x_1, \ldots, x_r$ are elements of $A(\Psi_0)$, then the product of $x_1, \ldots, x_r$ in $A$, with any grouping, is an element of $A(\Psi_0(r))$, as in the preceding paragraph. If

\[(22.5.10) \quad \Psi_0(r) \cap \Psi = \emptyset,\]

then such a product of $x_1, \ldots, x_r$ in $A$ is equal to 0.

If $y \in A = A(\Psi)$, then the product of $x_1, \ldots, x_r, y$ in $A$, with any ordering or grouping, is an element of $A(\Psi_0(r) + \Psi)$, as before. If

\[(22.5.11) \quad (\Psi_0(r) + \Psi) \cap \Psi = \emptyset,\]

then such a product of $x_1, \ldots, x_r, y$ in $A$ is equal to 0.

Suppose now that $A$ is a Lie algebra over $k$, so that (22.5.2) can be expressed as

\[(22.5.12) \quad [A(\Psi_1), A(\Psi_2)] \subseteq A(\Psi_1 + \Psi_2)\]

for every $\Psi_1, \Psi_2 \subseteq \Xi$. Equivalently, if $x \in \Psi_1$, then

\[(22.5.13) \quad \text{ad}_x(A(\Psi_2)) \subseteq A(\Psi_1 + \Psi_2).\]

If $\Psi_0 \subseteq \Xi, x \in A(\Psi_0)$, and $r \in \mathbb{Z}_+$, then

\[(22.5.14) \quad \text{ad}_x(A(\Psi_0(r) + \Psi)) \subseteq A(\Psi_0(r + 1) + \Psi).\]

Thus

\[(22.5.15) \quad (\text{ad}_x)^r(A) \subseteq A(\Psi_0(r) + \Psi)\]

for every $r \geq 1$. In particular, if $x \in A(\Psi_0)$ and $r \in \mathbb{Z}_+$ satisfies (22.5.11), then we get that $(\text{ad}_x)^r = 0$ on $A$.

Suppose that $\Psi_0$ satisfies (22.5.3), so that $A(\Psi_0)$ is a Lie subalgebra of $A$. Let $(A(\Psi_0))^j, j \geq 0$, be the lower central series in $A(\Psi_0)$, as in Section 9.5. If $r \in \mathbb{Z}_+$, then

\[(22.5.16) \quad (A(\Psi_0))^r - 1 \subseteq A(\Psi_0(r)),\]

as before. If (22.5.10) holds, then

\[(22.5.17) \quad (A(\Psi_0))^r - 1 = \{0\},\]

so that $A(\Psi_0)$ is nilpotent as a Lie algebra.

Suppose that $\Xi$ has an additive identity element 0, so that $A_0$ is a Lie subalgebra of $A$. If $\Psi_1, \Psi_2 \subseteq \Xi$, then

\[(22.5.18) \quad (\Psi_1 \cup \{0\}) + (\Psi_2 \cup \{0\}) = (\Psi_1 + \Psi_2) \cup \Psi_1 \cup \Psi_2 \cup \{0\}.\]

If $A_0$ is commutative as a Lie algebra, then

\[(22.5.19) \quad [A(\Psi_1 \cup \{0\}), A(\Psi_2 \cup \{0\})] \subseteq A((\Psi_1 + \Psi_2) \cup \Psi_1 \cup \Psi_2).\]
Suppose that \( \Psi_0 \subseteq \Xi \) satisfies (22.5.3) again. This implies that \( \Psi_0 \cup \{0\} \) satisfies the analogous condition,

\[
(22.5.20) \quad ((\Psi_0 \cup \{0\}) + (\Psi_0 \cup \{0\})) \cap \Psi \subseteq \Psi_0 \cup \{0\}.
\]

Thus \( A(\Psi_0 \cup \{0\}) \) is a Lie subalgebra of \( A \) too, and it is easy to see that \( A(\Psi_0) \) is an ideal in \( A(\Psi_0 \cup \{0\}) \). If \( A_0 \) is commutative as a Lie algebra, then

\[
(22.5.21) \quad [A(\Psi_0 \cup \{0\}), A(\Psi_0 \cup \{0\})] \subseteq A(\Psi_0).
\]

It follows that \( A(\Psi_0 \cup \{0\}) \) is solvable as a Lie algebra over \( k \) when (22.5.10) holds for some \( r \), so that \( A(\Psi_0) \) is nilpotent as a Lie algebra.

If

\[
(22.5.22) \quad [A_0, A_\alpha] = A_\alpha
\]

for every \( \alpha \in \Psi_0 \), then

\[
(22.5.23) \quad [A_0, A(\Psi_0)] = A(\Psi_0).
\]

If \( A_0 \) is commutative as a Lie algebra, then we get that

\[
(22.5.24) \quad [A(\Psi_0 \cup \{0\}), A(\Psi_0 \cup \{0\})] = A(\Psi_0).
\]

### 22.6 Roots and submodules

Let \( V \) be a vector space over the real numbers of positive finite dimension, and let \( \Phi \) be a root system in \( V \). Also let \( \Theta \) be the subgroup of \( V \), as a commutative group with respect to addition, generated by \( \Phi \). We would like to consider situations like those in Section 22.3, with

\[
(22.6.1) \quad \Xi = \Theta, \quad \Psi = \Phi \cup \{0\}.
\]

Note that \( \Theta \) has no nontrivial torsion, because \( V \) has none.

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( (A, [\cdot, \cdot]_A) \) be a Lie algebra over \( k \). If \( \alpha \in \Phi \cup \{0\} \), then we ask that \( A_\alpha \) be a submodule of \( A \), as a module over \( k \), and that \( A \) correspond to the direct sum of the \( A_\alpha \)'s, \( \alpha \in \Phi \cup \{0\} \), as a module over \( k \). As before, we put \( A_\alpha = \{0\} \) when \( \alpha \in \Theta \setminus (\Phi \cup \{0\}) \), and we ask that

\[
(22.6.2) \quad [A_\alpha, A_\beta] \subseteq A_{\alpha + \beta}
\]

for every \( \alpha, \beta \in \Theta \), as in (22.3.2). Thus \( A_0 \) is a Lie subalgebra of \( A \), as before. If \( x \in A_\alpha \) for some nonzero \( \alpha \in \Theta \), then \( x \) is ad-nilpotent in \( A \), as in Section 22.3.

If \( \Psi_0 \subseteq \Theta \), then let \( A(\Psi_0) \) be the subset of \( A \) consisting of finite sums of elements of \( A_\alpha, \alpha \in \Psi_0 \), as in Section 22.4. Let \( \Delta \subseteq \Phi \) be a base for \( \Phi \), as a root system in \( V \), and let \( \Phi^+ = \Phi^{\Delta^+} \) be the set of positive roots in \( \Phi \) with respect to \( \Delta \). Thus \( \Phi^+ \) consists of the \( \alpha \in \Phi \) that can be expressed as a linear
combination of elements of $\Delta$ whose coefficients are nonnegative integers, and 
$\Phi = \Phi^+ \cup (-\Phi^+)$. Note that 

\[(\Phi^+ + \Phi^+) \cap (\Phi \cup \{0\}) \subseteq \Phi^+,\]

so that $A(\Phi^+)$ is a Lie subalgebra of $A$, as before. Similarly,

\[((-\Phi^+) + (-\Phi^+)) \cap (\Phi \cup \{0\}) \subseteq -\Phi^+,\]

and $A(-\Phi^+)$ is a Lie subalgebra of $A$ too.

It is easy to see that

\[(\sum_{j=1}^{r} \Phi^+) \cap (\Phi \cup \{0\}) = \emptyset\]

when $r$ is sufficiently large, because $\Phi$ has only finitely many elements. In fact,

\[\left(\left(\sum_{j=1}^{r} \Phi^+\right) + (\Phi \cup \{0\})\right) \cap (\Phi \cup \{0\}) = \emptyset\]

when $r$ is sufficiently large. This implies that $A(\Phi^+)$ is nilpotent as a Lie algebra
over $k$, and that the elements of $A(\Phi^+)$ are ad-nilpotent in $A$, as in the previous
section. Similarly, $A(-\Phi^+)$ is nilpotent as a Lie algebra over $k$, and the elements
of $A(-\Phi^+)$ are ad-nilpotent in $A$. This corresponds to some remarks near the
top of p84 of [13], and part (b) of Theorem 4 on p47 of [23].

Observe that

\[((\Phi^+ \cup \{0\}) + (\Phi^+ \cup \{0\})) \cap (\Phi \cup \{0\}) \subseteq \Phi^+ \cup \{0\}\]

and

\[((-\Phi^+) \cup \{0\}) + ((-\Phi^+) \cup \{0\}) \cap (\Phi \cup \{0\}) \subseteq (-\Phi^+) \cup \{0\},\]

so that $A(\Phi^+ \cup \{0\})$ and $A((-\Phi^+) \cup \{0\})$ are Lie subalgebras of $A$. As in the
previous section, $A(\Phi^+)$ is an ideal in $A(\Phi^+ \cup \{0\})$, and $A(-\Phi^+)$ is an ideal in
$A((-\Phi^+) \cup \{0\})$.

If $A_0$ is commutative as a Lie algebra over $k$, then

\[[A(\Phi^+ \cup \{0\}), A(\Phi^+ \cup \{0\})] \subseteq A(\Phi^+)\]

and

\[[A((-\Phi^+) \cup \{0\}), A((-\Phi^+) \cup \{0\})] \subseteq A(-\Phi^+),\]

as in (22.5.21). This implies that $A(\Phi^+ \cup \{0\})$ and $A((-\Phi^+) \cup \{0\})$ are solvable as
Lie algebras over $k$, because $A(\Phi^+)$ and $A(-\Phi^+)$ are nilpotent as Lie algebras
over $k$. This corresponds to some of the remarks near the top of p84 of [13] again, and part of part (c) of Theorem 4 on p47 of [23].

If

\[[A_0, A_\alpha] = A_\alpha\]
for every \( \alpha \in \Phi \), then
\[
(22.6.12) \quad [A_0, A(\Phi^+)] = A(\Phi^+)
\]
and
\[
(22.6.13) \quad [A_0, A(-\Phi^+)] = A(-\Phi^+).
\]
If \( A_0 \) is commutative as a Lie algebra over \( k \), then it follows that
\[
(22.6.14) \quad [A(\Phi^+ \cup \{0\}), A(\Phi^+ \cup \{0\})] = A(\Phi^+)
\]
and
\[
(22.6.15) \quad [A((-\Phi^+) \cup \{0\}), A((-\Phi^+) \cup \{0\})] = A(-\Phi^+),
\]
as in (22.5.24). This corresponds to the second part of part (c) of Theorem 4 on p47 of [23].

22.7 Borel subalgebras

Let \( k \) be a commutative ring with a multiplicative identity element, and let
\((A, [\cdot, \cdot])_A\) be a Lie algebra over \( k \). A Lie subalgebra \( B \) of \( A \) is said to be a Borel subalgebra if \( B \) is solvable as a Lie algebra over \( k \), and \( B \) is maximal with respect to inclusion among Lie subalgebras of \( A \) that are solvable as Lie algebras. If solvable subalgebras in \( A \) satisfy an ascending chain condition, then every solvable subalgebra of \( A \) is contained in a Borel subalgebra. In particular, this holds for finite-dimensional Lie algebras over fields.

Let \( A \) be an algebra over \( k \) in the strict sense for the moment, where multiplication of \( a, b \in A \) is expressed as \( ab \), and let \( A_1, A_2 \) be subalgebras of \( A \). In particular, \( A_1 \) and \( A_2 \) are submodules of \( A \), as a module over \( k \), so that \( A_1 + A_2 \) is a submodule of \( A \). Let \( a_1, b_1 \in A_1 \) and \( a_2, b_2 \in A_2 \) be given, so that
\[
(22.7.1) \quad (a_1 + a_2)(b_1 + b_2) = a_1b_1 + a_1b_2 + a_2b_1 + a_2b_2.
\]
Of course, \( a_1b_1 \in A_1 \) and \( a_2b_2 \in A_2 \), by hypothesis. If \( A_1 \) is a right ideal in \( A \), then \( a_1b_2 \in A_1 \), and if \( A_2 \) is a left ideal in \( A \), then \( a_1b_2 \in A_2 \). Similarly, if \( A_1 \) is a left ideal in \( A \), then \( a_2b_1 \in A_1 \), and if \( A_2 \) is a right ideal in \( A \), then \( a_2b_1 \in A_2 \). It follows that (22.7.1) is an element of \( A_1 + A_2 \) when either \( A_1 \) is a right ideal in \( A \) or \( A_2 \) is a left ideal in \( A \), and either \( A_1 \) is a left ideal in \( A \), or \( A_2 \) is a right ideal in \( A \). In particular, this holds when \( A_1 \) or \( A_2 \) is a two-sided ideal in \( A \). This means that \( A_1 + A_2 \) is a subalgebra of \( A \) under these conditions.

Let \( A \) be a Lie algebra over \( k \) again, let \( A_1 \) be a Lie subalgebra of \( A \), and let \( A_2 \) be an ideal in \( A \). Thus \( A_1 + A_2 \) is a Lie subalgebra of \( A \), as in the preceding paragraph. If \( A_1 \) and \( A_2 \) are both solvable as Lie algebras over \( k \), then \( A_1 + A_2 \) is solvable as well. This was mentioned in Section 9.4 when \( A_1 \) and \( A_2 \) are both ideals in \( A \), and essentially the same argument works in this situation. More precisely, note that \( A_2 \) is an ideal in \( A_1 + A_2 \). It suffices to verify that \((A_1 + A_2)/A_2\) is solvable as a Lie algebra, because \( A_2 \) is solvable, as in Section 9.3. Let \( q_2 \) be the canonical quotient mapping from \( A_1 + A_2 \) onto
(A_1 + A_2)/A_2, which is essentially the same as the restriction of the canonical quotient mapping from \( A \) onto \( A/A_2 \) to \( A_1 + A_2 \).

It is easy to see that \( A_1 \cap A_2 \) is an ideal in \( A_1 \), because \( A_2 \) is an ideal in \( A \). In fact, \( A_1 \cap A_2 \) is the same as the kernel of the restriction of \( q_2 \) to \( A_1 \), as a subalgebra of \( A_1 + A_2 \). Note that \( q_2(A_1) = q_2(A_1 + A_2) \). The restriction of \( q_2 \) to \( A_1 \) can be identified with the canonical quotient mapping from \( A_1 \) onto \( A_1/(A_1 \cap A_2) \), as a Lie algebra homomorphism. This leads to a natural Lie algebra isomorphism between \( A_1/(A_1 \cap A_2) \) and \( (A_1 + A_2)/A_2 \). By hypothesis, \( A_1 \) is solvable, which implies that \( A_1/(A_1 \cap A_2) \) is solvable, as in Section 9.3. This means that \( (A_1 + A_2)/A_2 \) is solvable, as desired.

If \( A_1 \) is a Borel subalgebra of \( A \), then we get that \( A_1 = A_1 + A_2 \), which means that

\[
(22.7.2) \quad A_2 \subseteq A_1.
\]

If \( A \) has a solvable radical \( \text{Rad} \ A \), as in Section 9.4, then it follows that

\[
(22.7.3) \quad \text{Rad} \ A \subseteq A_1.
\]

This corresponds to part of the proof of Lemma B on p83 of [13].

Let \( A_0 \) be a solvable ideal in \( A \), and let \( q_0 \) be the canonical quotient mapping from \( A \) onto \( A/A_0 \). If \( B \) is a Lie subalgebra of \( A \) that is solvable as a Lie algebra over \( k \), then \( q_0(B) \) is solvable too, as in Section 9.3. If \( B_0 \) is a solvable Lie subalgebra of \( A/A_0 \), then \( q_0^{-1}(B_0) \) is a solvable Lie subalgebra of \( A \), as before. If \( B \) is a Borel subalgebra of \( A \), then \( A_0 \subseteq B \), as in the preceding paragraph. This leads to a natural one-to-one correspondence between Borel subalgebras of \( A \) and \( A/A_0 \), as in Lemma B on p83 of [13].

Let \( B \) be a Lie subalgebra of \( A \), and let \( x \) be an element of the normalizer \( N_A(B) \) of \( B \) in \( A \). Under these conditions, it is easy to see that

\[
(22.7.4) \quad B(x) = \{ t x + y : t \in k, y \in B \}
\]

is a Lie subalgebra of \( A \), with

\[
(22.7.5) \quad [B(x), B(x)] \subseteq B.
\]

If \( B \) is solvable, then it follows that \( B(x) \) is solvable as well. If \( B \) is a Borel subalgebra of \( A \), then we get that \( B = B(x) \), so that \( x \in B \). This means that \( N_A(B) = B \) when \( B \) is a Borel subalgebra of \( A \), as in Lemma A on p83 of [13].

### 22.8 Some diagonalizability conditions

Let \( \Xi \) be a commutative semigroup, with the semigroup operation expressed additively, and with an additive identity element 0. Also let \( \Phi \) be a set of nonzero elements of \( \Xi \). We would like to consider the same type of situation as in Section 22.3, with \( \Psi = \Phi \cup \{0\} \).

Let \( k \) be a commutative ring with a multiplicative identity element again, and let \((A, [\cdot, \cdot]_A)\) be a Lie algebra over \( k \). As before, suppose that \( A_0 \) is a submodule
of $A$, as a module over $k$, for each $\alpha \in \Phi \cup \{0\}$, and that $A$ corresponds to the direct sum of the $A_\alpha$'s, as a module over $k$. We put $A_\alpha = \{0\}$ when $\alpha \in \Xi \setminus (\Phi \cup \{0\})$, and ask that

$$\begin{equation} \tag{22.8.1} [A_\alpha, A_\beta] \subseteq A_{\alpha + \beta} \end{equation}$$

for every $\alpha, \beta \in \Xi$, as in (22.3.2). In particular, this means that $A_0$ is a Lie subalgebra of $A$.

Let $\text{Hom}_k(A_0, k)$ be the space of homomorphisms from $A_0$ into $k$, as modules over $k$. We put $A_0 = f\_{\Phi}g$, and ask that

$$\begin{equation} \tag{22.8.2} \text{ad}_w(x) = [w, x]_A = \phi_\alpha(w) x \end{equation}$$

In particular, this is equal to 0 when $\alpha = 0$. This means that $A_0$ is commutative as a Lie algebra. If $\alpha \in \Xi \setminus (\Phi \cup \{0\})$ and $x \in A_\alpha$, then $x = 0$, so that (22.8.5) holds trivially.

Note that (22.8.1) follows from (22.8.5) when $\alpha$ or $\beta$ is 0. If $w \in A_0$, and $x \in A_\alpha, \ y \in A_\beta$ for some $\alpha, \beta \in \Xi$, then

$$\begin{equation} \tag{22.8.6} [w, [x, y]_A]_A = [w, x]_A y + [x, w]_A y = \phi_\alpha(w) x + \phi_\beta(w) [x, y]_A = \phi_{\alpha + \beta}(w) [x, y]_A. \end{equation}$$

Of course, this is compatible with (22.8.1) and (22.8.5).

If $w \in A_0$ satisfies

$$\begin{equation} \tag{22.8.7} \phi_\alpha(w) = 0 \end{equation}$$

for every $\alpha \in \Phi$, then $w$ is in the center $Z(A)$ of $A$ as a Lie algebra. In particular, if $Z(A) = \{0\}$, then $w = 0$.

If $\alpha \in \Xi$, then $[A_0, A_\alpha] \subseteq A_\alpha$, as in (22.8.1). If there is a $w \in A_0$ such that $\phi_\alpha(w)$ has a multiplicative inverse in $k$, then we get that

$$\begin{equation} \tag{22.8.8} [A_0, A_\alpha] = A_\alpha. \end{equation}$$
22.9 Diagonalizability and subspaces

Let $k$ be a field, and let $(A, [\cdot, \cdot]_A)$ be a finite-dimensional Lie algebra over $k$. Suppose that $B$ is a Lie subalgebra of $A$ such that every element of $B$ is ad-diagonalizable as an element of $A$. This implies that $B$ is commutative as a Lie algebra, as in Section 17.1. Let $B'$ be the dual of $B$, as a vector space over $k$, and for each $\alpha \in B'$, let $A_\alpha$ be the set of $x \in A$ such that $\text{ad}_w(x) = \alpha(w) x$ for every $w \in B$. Consider the set $\Phi_B$ of $\alpha \in B'$ such that $\alpha \neq 0$ and $A_\alpha \neq \{0\}$.

Under these conditions, $\Phi_B$ has only finitely many elements, and $A$ corresponds to the direct sum of the subspaces $A_{\alpha}$ with $\alpha \in \Phi_B \cup \{0\}$, as a vector space over $k$, as in Section 17.2. Remember that $[A_\alpha, A_\beta] \subseteq A_{\alpha + \beta}$ for every $\alpha, \beta \in B'$, as before. In this situation, we can take $\Xi_B$ to be a subsemigroup of $B'$, as a commutative group with respect to addition, with

\begin{equation}
\Phi_B \cup \{0\} \subseteq \Xi_B.
\end{equation}

Note that

\begin{equation}
[B, A_\alpha] = A_\alpha
\end{equation}

for every $\alpha \in \Phi_B$. It follows that $[A_0, A_\alpha] = A_\alpha$ for every $\alpha \in \Phi_B$, because $B \subseteq A_0$, by the commutativity of $B$ as a Lie algebra.

As before, $A_0$ is the centralizer $C_A(B)$ of $B$ in $A$. Suppose that $C_A(B) = B$, so that $A_0 = B$, and

\begin{equation}
\text{Hom}_k(A_0, k) = A_0' = B'.
\end{equation}

In this case, we can take $\phi$ to be the obvious inclusion mapping of $\Xi_B$ into $\Phi_B$, so that (22.9.5) follows from the definition of $A_\alpha$.

Let us return for the moment to the situation before the preceding paragraph, so that we do not necessarily ask the centralizer of $B$ in $A$ to be $B$. Let $C$ be a linear subspace of $A$ such that

\begin{equation}
[B, C] \subseteq C.
\end{equation}

In particular, this holds when $C$ is a Lie subalgebra of $A$ that contains $B$. If $w \in B$, then

\begin{equation}
\text{ad}_w(C) \subseteq C,
\end{equation}

and $\text{ad}_w$ is diagonalizable on $A$, by hypothesis. This implies that the restriction of $\text{ad}_w$ to $C$ is diagonalizable on $C$, as in Section 10.6.

If $u, v \in B$, then $[u, v]_A = 0$, which implies that $\text{ad}_u$ and $\text{ad}_v$ commute on $A$, as in Section 2.4. Thus the restrictions of $\text{ad}_u$ and $\text{ad}_v$ to $C$ commute as well. This means that the restrictions of $\text{ad}_u$, $u \in B$, to $C$ are simultaneously diagonalizable on $C$, by standard arguments. Put

\begin{equation}
\Psi_C = \{ \alpha \in B': A_\alpha \cap C \neq \{0\} \},
\end{equation}

which is contained in $\Phi_B \cup \{0\}$. It follows that $C$ corresponds to the direct sum of $A_\alpha \cap C$ with $\alpha \in \Psi_C$, as a vector space over $k$. If $C$ is a Lie subalgebra of $A$, then

\begin{equation}
[A_\alpha \cap C, A_\beta \cap C] \subseteq A_{\alpha + \beta} \cap C.
\end{equation}
22.10 Injectivity on $\Psi$

Let us consider some situations as in Section 22.8, with $k$ a field. Thus we let $\Xi$ be a commutative semigroup, with the semigroup operation expressed additively, and with an additive identity element 0. We also let $\Phi$ be a finite set of nonzero elements in $\Xi$, and start with situations as in Section 22.3, with $\Phi = \Phi \cup \{0\}$.

Let $(A, [-, -])$ be a finite-dimensional Lie algebra over $k$, and let $A_0$ be a linear subspace of $A$, as a vector space over $k$, for each $f_0, g$. Suppose that $A_\alpha$ corresponds to the direct sum of the $A_\alpha$’s, $\alpha \in \Phi \cup \{0\}$, and put $A_\alpha = \{0\}$ when $\alpha \in \Xi \setminus (\Phi \cup \{0\})$. As before, we ask that $[A_\alpha, A_\beta] \subseteq A_{\alpha+\beta}$ for every $\alpha, \beta \in \Xi$, which implies that $A_0$ is a Lie subalgebra of $A$.

Suppose that

$$\phi$$

is a semigroup homomorphism from $\Xi$ into $A'_0$,

where $A'_0$ is the dual of $A_0$ as a vector space over $k$, considered as a commutative semigroup with respect to addition in (22.10.1). As before, $\phi(0) = 0$ in $A'_0$, because $A'_0$ is a group with respect to addition. If $\alpha \in \Xi$, then it is convenient to put $\phi_\alpha = \phi(\alpha)$ again. We ask that $\text{ad}_w(x) = [w, x]_A = \phi_\alpha(w)x$ for every $\alpha \in \Phi \cup \{0\}$, $w \in A_0$, and $x \in A_\alpha$, which holds trivially when $\alpha \in \Xi \setminus (\Phi \cup \{0\})$. This implies that $A_0$ is commutative as a Lie subalgebra of $A$.

Note that $[A_0, A_\alpha] = A_\alpha$ for every $\alpha \in \Phi$ such that $\phi_\alpha \neq 0$, as in (22.8.8). More precisely, if $\phi_\alpha \neq 0$, then there is a $w \in A_0$ such that $\phi_\alpha(w) \neq 0$, and $\text{ad}_w(A_\alpha) = A_\alpha$.

Let us take $B = A_0$, so that every element of $B$ is ad-diagonalizable, by hypothesis. If $\rho$ is an element of the dual $B' = A'_0$ of $B$, then put

$$\tilde{A}_\rho = \{x \in A : \text{ for every } w \in B, \text{ ad}_w(x) = [w, x]_A = \rho(w)x\},$$

which is a linear subspace of $A$. Thus

$$A_\alpha \subseteq \tilde{A}_{\phi_\alpha}$$

for every $\alpha \in \Xi$. If $\rho$ is any element of $B'$, then it follows that $\tilde{A}_\rho$ corresponds to the direct sum of $A_\alpha$ with $\alpha \in \Phi \cup \{0\}$ and $\phi_\alpha = \rho$. This means that $\tilde{A}_\rho = \{0\}$ when $\rho \neq \phi_\alpha$ for every $\alpha \in \Phi \cup \{0\}$.

Suppose that

$$\text{ for every } \alpha \in \Phi, \text{ we have that } \phi_\alpha \neq 0.$$
This implies that

\[(22.10.5)\]

\[A_0 = \tilde{A}_0,\]

as in the preceding paragraph. Note that \(\tilde{A}_0\) is the same as the centralizer of \(B\) in \(A\), as before. Thus the centralizer \(C_A(A_0)\) of \(A_0\) in \(A\) is equal to \(A_0\) in this case. We also get that \([A_0, A_0] = A_0\) for every \(\alpha \in \Phi\) when (22.10.4) holds.

Suppose now that

\[(22.10.6)\]

\(\phi\) is injective on \(\Phi \cup \{0\}\),

so that \(\phi_\alpha \neq \phi_\beta\), as elements of \(A'_0\), for every \(\alpha, \beta \in \Phi \cup \{0\}\) with \(\alpha \neq \beta\). Of course, (22.10.6) is the same as saying that (22.10.4) holds, and that \(\phi\) is injective on \(\Phi\). It follows that

\[(22.10.7)\]

\[A_\alpha = \tilde{A}_{\phi_\alpha}\]

for every \(\alpha \in \Phi \cup \{0\}\), as before.

Let \(C\) be a linear subspace of \(A\) such that

\[(22.10.8)\]

\([A_0, C] \subseteq C,\]

and put

\[(22.10.9)\]

\[\tilde{\Psi}_C = \{\rho \in B' : \tilde{A}_\rho \cap C \neq \{0\}\},\]

which corresponds to (22.9.6) with the notation in this section. As in the previous section, \(C\) corresponds to the direct sum of \(\tilde{A}_\rho \cap C\) with \(\rho \in \tilde{\Psi}_C\), as a vector space over \(k\). Note that this does not use (22.10.4) or (22.10.6). Put

\[(22.10.10)\]

\[\Psi_C = \{\alpha \in \Phi \cup \{0\} : A_\alpha \cap C \neq \{0\}\}.\]

If (22.10.6) holds, then \(C\) corresponds to the direct sum of \(A_\alpha \cap C\) with \(\alpha \in \Psi_C\), as a vector space over \(k\).

### 22.11 Some additional conditions

Let \(V\) be a vector space over the real numbers of positive finite dimension, and let \(\Phi\) be a root system in \(V\). Thus, for every \(\alpha \in \Phi\), there is a unique symmetry \(\sigma_\alpha\) on \(V\) with vector \(\alpha\) that maps \(\Phi\) onto itself. Let \(\lambda_\alpha\) be the linear functional on \(V\) that is equal to 0 on the hyperplane fixed by \(\sigma_\alpha\) and satisfies \(\lambda_\alpha(\alpha) = 2\), as before. Equivalently, \(\sigma_\alpha(v) = v - \lambda_\alpha(v) \alpha\) for every \(v \in V\). Remember that \(\lambda_\alpha(\beta) \in \mathbb{Z}\) for every \(\alpha, \beta \in \Phi\), by definition of a root system.

Let \(\Theta\) be the subgroup of \(V\), as a commutative group with respect to addition, generated by \(\Phi\). This is the same as the subsemigroup of \(V\) generated by \(\Phi\). We would like to consider situations like those discussed in Section 22.8, with \(\Xi = \Theta\). Note that the elements of \(\Theta\) can be expressed as linear combinations of the elements of any base for \(\Phi\) with coefficients in \(\mathbb{Z}\).

Let \(k\) be a commutative ring with a multiplicative identity element, and let \((A, [\cdot, \cdot], A)\) be a Lie algebra over \(k\). If \(\alpha \in \Phi \cup \{0\}\), then we suppose that \(A_\alpha\) is a submodule of \(A\), as a module over \(k\), and that \(A\) corresponds to the direct sum of the \(A_\alpha\)'s, as a module over \(k\). As usual, it is convenient to put \(A_\alpha = \{0\}\).
when \( \alpha \in \Theta \setminus (\Phi \cup \{0\}) \), and we ask that \([A_\alpha, A_\beta] \subseteq A_{\alpha + \beta}\) for every \(\alpha, \beta \in \Theta\).

We also ask that \(\phi\) be a group homomorphism from \(\Theta\) into \(\text{Hom}_k(A_0, k)\), as a commutative group with respect to addition, such that \(\text{ad}_u(x) = \phi_\alpha(x)\) for every \(u \in A_0\) and \(x \in A_\alpha, \alpha \in \Phi \cup \{0\}\), where \(\phi_\alpha = \phi(\alpha)\). This implies that \(A_0\) is a Lie subalgebra of \(A\) that is commutative as a Lie algebra over \(k\), as before.

If \(\alpha \in \Phi\), then \(-\alpha \in \Phi\), and \([A_\alpha, A_{-\alpha}] \subseteq A_0\). We may be interested in situations where there is an \(h_\alpha \in A_0\) such that

\[
[A_\alpha, A_{-\alpha}] = \{ t h_\alpha : t \in k \}.
\]

We may also ask that

\[
\phi_\alpha(h_\alpha) = 2(= 1 + 1)
\]

in \(k\). In particular, (22.11.1) implies that there are \(x_\alpha \in A_\alpha\) and \(y_\alpha \in A_{-\alpha}\) such that

\[
[x_\alpha, y_\alpha]_A = h_\alpha.
\]

More precisely, we may ask that

\[
\phi_\beta(h_\alpha) = \lambda_\alpha(\beta) \cdot 1
\]

in \(k\), which reduces to (22.11.2) when \(\alpha = \beta\).

In the situation considered in Section 22.1, we take \(V = E_R\) and \(\Phi = \Phi_B\). More precisely, \(k\) is a field of characteristic 0 in this case, and \(\Phi\) may be considered as a subset of the dual of \(A_0 = B\), as a vector space over \(k\).

Remember that \(E_Q\) is the linear subspace of the dual of \(A_0\), as a vector space over \(Q\), spanned by \(\Phi\). By construction, \(E_Q\) is also a linear subspace of \(V = E_R\), as a vector space over \(Q\). In particular,

\[
\Theta \subseteq E_Q.
\]

Thus \(\phi\) may be taken to be the obvious inclusion mapping of \(\Theta\) into the dual of \(A_0\). The one-dimensionality of \([A_\alpha, A_{-\alpha}]\) for \(\alpha \in \Phi\) was discussed in Section 17.5, and \(h_\alpha\) was defined in Section 17.6. The one-dimensionality of \(A_\alpha, A_{-\alpha}\) was discussed in Section 17.7, and (22.11.5) follows from the definition of \(\lambda_\alpha\) in Section 22.1.

We also have that

\[
A_0 \text{ is spanned by the } h_\alpha's, \alpha \in \Phi,
\]

as a vector space over \(k\), in the situation considered in Section 22.1. More precisely, \(A_0 = B\) is spanned by \(t_{h_\alpha}, \alpha \in \Phi\), because the dual of \(B\) is spanned by \(\Phi\).
Similarly, if $\Delta \subseteq \Phi$ is a base for $\Phi$, then

$$A_0 \text{ is spanned by the } h_\alpha \text{'s, } \alpha \in \Delta, \quad (22.11.8)$$

as a vector space over $k$, in the situation considered in Section 22.1. To see this, it suffices to check that $A_0$ is spanned by $t_{b,\alpha}$, $\alpha \in \Delta$, because $h_\alpha$ was defined to be the product of $t_{b,\alpha}$ by a nonzero element of $k$. This reduces to the previous case, because every element of $\Phi$ can be expressed as a linear combination of elements of $\Delta$ with integer coefficients.

In fact, $\{ h_\alpha, \alpha \in \Delta \}$ form a basis for $A_0$ as a vector space over $k$, in the situation considered in Section 22.1. To see this, remember that the number of elements of $\Delta$ is the same as the dimension of $V = E_{\mathbb{R}}$ as a vector space over $\mathbb{R}$. This is the same as the dimension of $E_{\mathbb{Q}}$, as a vector space over $\mathbb{Q}$. This is also the same as the dimension of the dual of $A_0 = B$, as a vector space over $k$. Of course, this is the same as the dimension of $A_0$, as a vector space over $k$.

If $\alpha, \beta \in \Phi$, then $\text{ad}_{x_\alpha}$ maps $A_\beta$ into $A_{\alpha+\beta}$, as before. In the situation considered in Section 22.1, if we also have that $\alpha + \beta \in \Phi$, then

$$\text{ad}_{x_\alpha}(x_\beta) = \phi_{\alpha} x_\beta + \phi_{\beta} x_\alpha \quad (22.11.10)$$

and $\text{ad}_{x_\alpha}$ is a one-to-one mapping onto $A_{\alpha+\beta}$.

### 22.12 Standard Borel subalgebras

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $\Phi$ be a root system in $V$. Also let $\Theta$ be the subgroup of $V$, as a commutative group with respect to addition, generated by $\Phi$. We would like to consider situations with some of the conditions mentioned in the previous section, with $k$ a field. Let $(A, \{ [\cdot, \cdot]_A \})$ be a finite-dimensional Lie algebra over $k$, and let $A_\alpha$ be a linear subspace of $A$, as a vector space over $k$, for each $\alpha \in \Phi \cup \{0\}$. As before, we suppose that $A$ corresponds to the direct sum of the $A_\alpha$'s, $\alpha \in \Phi \cup \{0\}$, as a vector space over $k$, and we put $A_\alpha = \{0\}$ when $\alpha \in \Theta \setminus (\Phi \cup \{0\})$, for convenience. We suppose that $[A_\alpha, A_\beta] \subseteq A_{\alpha+\beta}$ for every $\alpha, \beta \in \Theta$, and that there is a group homomorphism $\phi$ from $\Theta$ into the dual $A_0$ of $A_0$ as a vector space over $k$ such that $\phi_{\alpha} x = \phi_{\alpha}(w)x$ for every $w \in A_0$ and $x \in A_\alpha$, $\alpha \in \Phi \cup \{0\}$. Here $\phi_{\alpha} = \phi(\alpha)$ for every $\alpha \in \Theta$, as usual.

Suppose that for each $\alpha \in \Phi$ there are $x_\alpha \in A_\alpha$ and $y_\alpha \in A_{-\alpha}$ such that $A_\alpha, A_{-\alpha}$ are spanned in $A$ by $x_\alpha, y_\alpha$, respectively. Put $h_\alpha = [x_\alpha, y_\alpha]_A$ for every $\alpha \in \Phi$, so that $h_\alpha \in A_0$. We ask that

$$\phi_{\alpha}(h_\alpha) = 2 \cdot 1 \quad (22.12.1)$$

in $k$ for every $\alpha \in \Phi$, as before. Suppose in addition that

$$\phi \text{ is injective on } \Phi \cup \{0\}, \quad (22.12.2)$$
22.13. SUBALGEBRAS AND DERIVATIONS

so that $\phi_\alpha \neq \phi_\beta$ when $\alpha$ and $\beta$ are distinct elements of $\Phi \cup \{0\}$. Note that this holds in the situation considered in Section 22.1, as in the previous section.

In particular, we are in the type of situation considered in Section 22.6. Let $\Delta \subseteq \Phi$ be a base for $\Phi$, as a root system in $V$, and let $\Phi^+$ be the set of positive roots in $\Phi$ with respect to $\Delta$, as before. If $\Psi_0 \subseteq \Theta$, then we let $A(\Psi_0)$ be the linear subspace of $A$ spanned by $A_\alpha$ with $\alpha \in \Psi_0$, as in Section 22.4. Remember that $A_0$ is commutative as a Lie subalgebra of $A$, as in Section 22.8. It follows that $A(\Phi^+ \cup \{0\})$ is a solvable Lie subalgebra of $A$, as in Section 22.6.

Let $C$ be a Lie subalgebra of $A$ such that

\[(22.12.3) \quad A(\Phi^+ \cup \{0\}) \subseteq C\]

In particular, this means that $A_0 \subseteq C$, so that $[A_0, C] \subseteq C$. If we put

\[(22.12.4) \quad \Psi_C = \{\alpha \in \Phi \cup \{0\} : A_\alpha \cap C \neq \{0\}\}, \]

then $C$ corresponds to the direct sum of $A_\alpha \cap C$ with $\alpha \in \Psi_C$, as in Section 22.10. Note that this uses (22.12.2).

Suppose that $A(\Phi^+ \cup \{0\})$ is a proper subset of $C$. This implies that $\Psi_C$ is not contained in $\Phi^+ \cup \{0\}$, as in the preceding paragraph. This means that there is an $\alpha \in \Psi_C \cap (-\Phi^+)$, because $\Phi = \Phi^+ \cup (-\Phi^+)$. It follows that $x_\alpha \in C$, because $A_\alpha$ is spanned by $x_\alpha$.

Under these conditions, $-\alpha \in \Phi^+$, and $y_\alpha, h_\alpha \in A(\Phi^+ \cup \{0\})$, so that $y_\alpha, h_\alpha$ are elements of $C$ too. Suppose that the characteristic of $k$ is not 2, which implies that

\[(22.12.5) \quad \phi_\alpha(h_\alpha) \neq 0.\]

In particular, $x_\alpha, y_\alpha, h_\alpha \neq 0$, and their linear span in $A$ is a Lie subalgebra of $A$ that is isomorphic to $sl_2(k)$, as a Lie algebra over $k$. Remember that this Lie subalgebra is not solvable in this case. It follows that $C$ is not solvable as a Lie algebra over $k$.

Equivalently, if the characteristic of $k$ is not 2, and if $C$ is solvable as a Lie algebra, then $A(\Phi^+ \cup \{0\}) = C$. This means that $A(\Phi^+ \cup \{0\})$ is a Borel subalgebra of $A$ when the characteristic of $k$ is not 2. This may be called the standard Borel subalgebra of $A$ with respect to $\Phi, \Delta$, and the $A_\alpha$’s. This corresponds to some remarks near the top of p84 in [13], and at the bottom of p47 of [23].

22.13 Subalgebras and derivations

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be an algebra over $k$ in the strict sense, where multiplication of $x, y \in A$ is expressed as $xy$. If $B$ and $C$ are submodules of $A$, as a module over $k$, then remember that their product $B \cdot C$ is defined to be the subset of $A$ consisting of finite sums of elements of $A$ of the form $xy$, with $x \in B$ and $y \in C$. This is a submodule of $A$ too, as a module over $k$, as in Section 9.2.
CHAPTER 22. ROOTS AND LIE ALGEBRAS

Let $B$ be a submodule of $A$ again, as a module over $k$, and put $B_1 = B$. If $B_1, \ldots, B_l \subseteq A$ have been defined for some positive integer $l$, then let $B_{l+1} \subseteq A$ be the set of finite sums of elements of $B_j \cdot B_{l+1-j}$, $j = 1, \ldots, l$. Continuing in this way, we can define $B_l$ as a submodule of $A$, as a module over $k$, for every $l \in \mathbb{Z}_+$. Equivalently, $B_l$ consists of finite sums of elements of $A$ that can be expressed as the product of $l$ elements of $B$, with some grouping of the factors.

Let $B_\infty \subseteq A$ be the set of finite sums of elements of $B_l$, $l \in \mathbb{Z}_+$. It is easy to see that $B_\infty$ is a submodule of $A$ as a module over $k$, and in fact a subalgebra of $A$. Of course, $B = B_1 \subseteq B_\infty$, by construction. More precisely, $B_\infty$ is the smallest subalgebra of $A$ that contains $B$. We may call $B_\infty$ the subalgebra of $A$ generated by $B$.

Let $\delta$ be a derivation on $A$ such that

\begin{equation}
\delta(B) \subseteq B.
\end{equation}

One can check that

\begin{equation}
\delta(B_l) \subseteq B_l
\end{equation}

for every $l \in \mathbb{Z}_+$, by induction. This implies that

\begin{equation}
\delta(B_\infty) \subseteq B_\infty.
\end{equation}

22.14 Reducibility and semigroups

Let us return to the type of situation considered in Section 22.3. Thus we let $\Xi$ be a commutative semigroup, with the semigroup operation expressed additively, and we let $\Psi$ be a nonempty subset of $\Xi$. Also let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [[\cdot,\cdot]_A)$ be a Lie algebra over $k$. Suppose that $A_\alpha$ is a submodule of $A$, as a module over $k$, for each $\alpha \in \Psi$, and that $A$ corresponds to the direct sum of the $A_\alpha$’s, $\alpha \in \Psi$, as a module over $k$. It is convenient to put $A_\alpha = \{0\}$ when $\alpha \in \Xi \setminus \Psi$, as before. We ask that $[A_\alpha, A_\beta] \subseteq A_{\alpha+\beta}$ for every $\alpha, \beta \in \Psi$, which holds trivially when $\alpha$ or $\beta$ is in $\Xi \setminus \Psi$. If $\Xi$ has an additive identity element 0, then it follows that $A_0$ is a Lie subalgebra of $A$.

If $\Psi_0 \subseteq \Xi$, then let $A(\Psi_0)$ be the subset of $A$ consisting of finite sums of elements of $A_\alpha$, $\alpha \in \Psi_0$, as in Section 22.4. If $(\Psi_0 + \Psi_0) \cap \Psi \subseteq \Psi_0$, then $A(\Psi_0)$ is a Lie subalgebra of $A$, as before.

Let $\Psi_1, \Psi_2$ be nonempty subsets of $\Psi$. If

\begin{equation}
(\Psi_j + \Psi_j) \cap \Psi \subseteq \Psi_j
\end{equation}

for $j = 1, 2$, then $A(\Psi_1), A(\Psi_2)$ are Lie subalgebras of $A$, as in the preceding paragraph. If

\begin{equation}
\Psi_1 \cup \Psi_2 = \Psi
\end{equation}

and

\begin{equation}
\Psi_1 \cap \Psi_2 = \emptyset,
\end{equation}

then $A(\Psi_1) \cap A(\Psi_2) = \{0\}$. If

\begin{equation}
\Psi_1 \subseteq \Psi_2
\end{equation}

then $A(\Psi_1) \subseteq A(\Psi_2)$.
then $A$ corresponds to the direct sum of $A(\Psi_1)$ and $A(\Psi_2)$, as a module over $k$.

If
\begin{equation}
(\Psi_1 + \Psi_2) \cap \Psi = \emptyset,
\end{equation}
then it follows that
\begin{equation}
[A(\Psi_1), A(\Psi_2)] = \{0\}.
\end{equation}

If all four of these conditions hold, then $A$ corresponds to the direct sum of $A(\Psi_1)$ and $A(\Psi_2)$, as a Lie algebra over $k$.

Suppose from now on in this section that $\Xi$ has an additive identity element 0, and that $0 \in \Psi$. In this case, (22.14.2) implies that 0 is an element of $\Psi_1$ or $\Psi_2$. This means that (22.14.4) does not hold, because $\Psi_1, \Psi_2 \neq \emptyset$.

Let $\Phi_1, \Phi_2$ be nonempty subsets of
\begin{equation}
\Phi = \Psi \setminus \{0\}.
\end{equation}

If $j \in \{1, 2\}$ and
\begin{equation}
(\Phi_j + \Phi_j) \cap \Psi \subseteq \Phi_j \cup \{0\},
\end{equation}
then
\begin{equation}
((\Phi_j \cup \{0\}) + (\Phi_j \cup \{0\})) \cap \Psi \subseteq \Phi_j \cup \{0\},
\end{equation}
and $A(\Phi_j \cup \{0\})$ is a Lie subalgebra of $A$, as before. Let
\begin{equation}
A^j
\end{equation}
be the Lie subalgebra of $A$ generated by $A(\Phi_j)$,

so that
\begin{equation}
A^j \subseteq A(\Phi_j \cup \{0\})
\end{equation}
in this situation.

If
\begin{equation}
(\Phi_1 + \Phi_2) \cap \Psi = \emptyset,
\end{equation}
then
\begin{equation}
[A(\Phi_1), A(\Phi_2)] = \{0\},
\end{equation}
as before. Equivalently, this means that
\begin{equation}
A(\Phi_1) \subseteq C_A(A(\Phi_2)), \quad A(\Phi_2) \subseteq C_A(A(\Phi_1)).
\end{equation}

This implies that
\begin{equation}
A^1 \subseteq C_A(A(\Phi_2)), \quad A^2 \subseteq C_A(A(\Phi_1)),
\end{equation}
because the centralizer $C_A(E)$ of any $E \subseteq A$ in $A$ is a Lie subalgebra of $A$, as in Section 7.6. This is the same as saying that
\begin{equation}
A(\Phi_2) \subseteq C_A(A^1), \quad A(\Phi_1) \subseteq C_A(A^2).
\end{equation}

It follows that
\begin{equation}
A^2 \subseteq C_A(A^1), \quad A^1 \subseteq C_A(A^2),
\end{equation}
as before, which are of course equivalent.

If \( w \in A_0 \), then
\[
(22.14.17) \quad \text{ad}_w(A(\Phi_j)) \subseteq A(\Phi_j)
\]
for \( j = 1, 2 \), because \( \text{ad}_w(A_\alpha) \subseteq A_\alpha \) for every \( \alpha \in \Psi \), by hypothesis. This
implies that
\[
(22.14.18) \quad \text{ad}_w(A^j) \subseteq A^j
\]
for \( j = 1, 2 \), as in the previous section, because \( \text{ad}_w \) is a derivation on \( A \).

Let us check that
\[
(22.14.19) \quad A^j \text{ is an ideal in } A \text{ for } j = 1, 2
\]
when (22.14.7) holds. Of course, this is the same as saying that the normalizer
\( N_A(A^j) \) of \( A^j \) in \( A \) is equal to \( A \). By construction, \( A(\Phi_j) \subseteq A^j \subseteq N_A(A^j) \) for
\( j = 1, 2 \). Using (22.14.14), we get that \( A(\Phi_2) \subseteq N_A(A^1) \) and \( A(\Phi_1) \subseteq N_A(A^2) \).
We also have that \( A_0 \subseteq N_A(A^j) \) for \( j = 1, 2 \), by (22.14.18), so that \( N_A(A^j) = A \),
as desired.

If
\[
(22.14.20) \quad \text{for each } j = 1, 2, \text{ there is an } \alpha_j \in \Phi_j \text{ such that } A_{\alpha_j} \neq \{0\},
\]
then
\[
(22.14.21) \quad A^j \neq \{0\} \text{ for } j = 1, 2.
\]
If (22.14.7) holds for \( j = 1, 2 \), (22.14.20) holds, and
\[
(22.14.22) \quad \Phi_1 \cap \Phi_2 = \emptyset,
\]
then
\[
(22.14.23) \quad A^j \neq A \text{ for } j = 1, 2,
\]
by (22.14.10). If (22.14.11) holds too, and \( k \) is a field, then it follows that \( A \)
is not simple as a Lie algebra, because of (22.14.19). This is related to the
proposition on p73 of [13], and to part of the corollary to Theorem 9 on p50 of
[23].

Note that
\[
(22.14.24) \quad A^1 \cap A^2 \subseteq A_0
\]
when (22.14.7) holds for \( j = 1, 2 \) and (22.14.22) holds, because of (22.14.10). If
(22.14.11) holds as well,
\[
(22.14.25) \quad \Phi_1 \cup \Phi_2 = \Phi,
\]
and
\[
(22.14.26) \quad A_0 \text{ is commutative as a Lie algebra over } k,
\]
than
\[
(22.14.27) \quad A^1 \cap A^2 \subseteq Z(A),
\]
where \( Z(A) \) is the center of \( A \) as a Lie algebra, as usual. More precisely, if
\( x \in A^1 \cap A^2 \), then \( [x, y]_A = 0 \) for every \( y \in A^1 \) or \( A^2 \), by (22.14.16). This holds
when \( y \in A_0 \) too, by (22.14.24) and (22.14.26). It follows that this holds for every \( y \in A \), because of (22.14.25), as desired.

If \( Z(A) = \{0\} \), then (22.14.27) implies that

\[
A^1 \cap A^2 = \{0\}. 
\]

(22.14.28)

If we also have that

\[
A \text{ every element of } A \text{ can be expressed as the sum of elements of } A^1 \text{ and } A^2, 
\]

then \( A \) corresponds to the direct sum of \( A^1 \) and \( A^2 \), as a module over \( k \). If (22.14.11) holds too, then \( A \) corresponds to the direct sum of \( A^1 \) and \( A^2 \) as a Lie algebra over \( k \), because of (22.14.16). If (22.14.25) holds, and every element of \( A_0 \) can be expressed as a sum of elements of \( A^1 \) and \( A^2 \), then (22.14.29) holds.

If \( j \in \{1, 2\} \), \( \alpha, \beta \in \Phi_j \), and \( \alpha + \beta = 0 \), then

\[
[A_\alpha, A_\beta] \subseteq A^j \cap A_0. 
\]

(22.14.30)

If (22.15) and (22.14.26) hold, then one can check that set of elements of \( A \) that can be expressed as finite sums of elements of \( A(\Phi_j) \) and \( [A_\alpha, A_\beta] \) with \( \alpha, \beta \in \Phi_j \) and \( \alpha + \beta = 0 \) is a Lie subalgebra of \( A \). It follows that this is the same as \( A^j \) in this case.

## 22.15 Reducibility and roots

Let \( V \) be a vector space over the real numbers of positive finite dimension, and let \( \Phi \) be a root system in \( V \). Also let \( \Theta \) be the subgroup of \( V \), as a group with respect to addition, generated by \( \Phi \). We would like to consider situations like those in the previous section, with \( \Xi = \Theta \) and \( \Psi = \Phi \cup \{0\} \). Thus we let \( k \), \( A \), and \( A_0 \) be as before.

Suppose that \( \Phi \) is reducible as a root system in \( V \), as in Section 20.4. Thus \( V \) is the direct sum of two nontrivial linear subspaces \( V_1 \) and \( V_2 \), as a vector space over \( \mathbb{R} \), and

\[
\Phi \subseteq V_1 \cup V_2. 
\]

(22.15.1)

Put \( \Phi_j = \Phi \cap V_j \) for \( j = 1, 2 \), so that \( \Phi = \Phi_1 \cup \Phi_2 \). If \( \alpha \in \Phi_1 \) and \( \beta \in \Phi_2 \), then \( \alpha \neq \beta \), because \( V_1 \cap V_2 = \{0\} \), and

\[
\alpha + \beta \not\in \Phi \cup \{0\}, 
\]

(22.15.2)

because \( \alpha + \beta \not\in V_1 \cup V_2 \). This means that (22.14.7), (22.14.11), (22.14.22), and (22.14.25) hold in this situation.

If \( j \in \{1, 2\} \) and \( \alpha \in \Phi_j \), then \( -\alpha \in \Phi_j \), and

\[
[A_\alpha, A_{-\alpha}] \subseteq A^j \cap A_0, 
\]

(22.15.3)

as in (22.14.30). Suppose from now on in this section that \( A_0 \) is commutative as a Lie algebra over \( k \), as in (22.14.26). This implies that the set of elements of
CHAPTER 22. ROOTS AND LIE ALGEBRAS

A that can be expressed as finite sums of elements of \( A(\Phi_j) \) and \([A_\alpha, A_{-\alpha}]\) with \( \alpha \in \Phi_j \) is a Lie subalgebra of \( A \), as in the previous section. This Lie subalgebra is the same as \( A^j \) under these conditions, as before.

We can describe \( A^j, j = 1, 2 \), more precisely with some additional conditions, as in Section 22.11. Suppose that for each \( \alpha \in \Phi \), there is an \( h_\alpha \in A_0 \) such that \([A_\alpha, A_{-\alpha}]\) is the set of multiples of \( h_\alpha \) by elements of \( k \), as before. This means that \( A^j \) consists of sums of elements of \( A(\Phi_j) \) and linear combinations of the \( h_\alpha \)'s, \( \alpha \in \Phi_j \), with coefficients in \( k \), for \( j = 1, 2 \).

Let \( \Delta_1 \) and \( \Delta_2 \) be bases for \( \Phi_1 \) and \( \Phi_2 \) in \( V_1 \) and \( V_2 \), respectively. This implies that \( \Delta_1 \cup \Delta_2 \) is a base for \( \Phi \), as in Section 20.4. In the situation considered in Section 22.1, \( k \) is a field, and \( h_\alpha, \alpha \in \Delta_1 \cup \Delta_2 \), form a basis for \( A_0 \) as a vector space over \( k \), as in Section 22.11. If \( \alpha \in \Phi_j \) for \( j = 1, 2 \), then \( h_\alpha \) can be expressed as a linear combination of \( h_\beta, \beta \in \Delta_j \), in \( A_0 \) in this case. This follows from the fact that \( \alpha \) can be expressed as a linear combination of the elements of \( \Delta_j \) with integer coefficients, so that \( t_{b,\alpha} \) can be expressed as a linear combination of \( t_{b,\beta}, \beta \in \Delta_j \), where \( t_{b,\gamma} \) is as in Section 22.1.

Under these conditions, \( A^j \) consists of sums of elements of \( A(\Phi_j) \) and linear combinations of the \( h_\alpha \)'s, now with \( \alpha \in \Delta_j \), for \( j = 1, 2 \). It follows that every element of \( A \) can be expressed in a unique way as a sum of elements of \( A^1 \) and \( A^2 \), so that \( A \) corresponds to the direct sum of \( A^1 \) and \( A^2 \), as a Lie algebra over \( k \).

As before, this is related to the proposition on p73 of [13], and to part of the corollary to Theorem 9 on p50 of [23].
Chapter 23

Roots and Lie algebras, 2

23.1 Inverse roots and homomorphisms

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $\Phi$ be a root system in $V$. If $\alpha \in \Phi$, then there is a unique symmetry $\sigma_\alpha$ on $V$ with vector $\alpha$ that maps $\Phi$ onto itself, as usual. Let $\lambda_\alpha$ be the linear functional on $V$ that is equal to 0 on the hyperplane fixed by $\sigma_\alpha$ and satisfies $\lambda_\alpha(\alpha) = 2$, so that $\sigma_\alpha(v) = v - \lambda_\alpha(v)\alpha$ for every $v \in V$. Remember that

$$\Phi' = \{\lambda_\alpha : \alpha \in \Phi\}$$

is a root system in the dual $V'$ of $V$, as in Section 19.8.

Note that

$$\alpha \mapsto \lambda_\alpha$$

is a one-to-one mapping from $\Phi$ onto $\Phi'$. More precisely, if $(\cdot, \cdot)$ is an inner product on $V$ that is invariant under the Weyl group of $\Phi$ and $\alpha \in \Phi$, then $\lambda_\alpha$ corresponds to $\hat{\alpha} = 2 \alpha (\alpha, \alpha)^{-1}$ with respect to $(\cdot, \cdot)$, and $\hat{\alpha} = \alpha$. Alternatively, the symmetry on $V'$ with vector $\lambda_\alpha$ that maps $\Phi'$ onto itself corresponds to the dual of $\sigma_\alpha$, and the associated linear functional on $V'$ corresponds to $\alpha$ in a natural way.

Let $\Theta = \Theta_\Delta$ be the subgroup of $V$, as a commutative group with respect to addition, generated by $\Phi$. If $\Delta$ is a base for $\Phi$, then $\Theta$ consists of the linear combinations of elements of $\Delta$ with coefficients in $\mathbb{Z}$.

Let $\text{Hom}(\Theta, \mathbb{Z})$ be the collection of homomorphisms from $\Theta$ in $\mathbb{Z}$, as commutative groups with respect to addition. Of course, commutative groups may be considered as modules over $\mathbb{Z}$, and group homomorphisms between them may be considered as module homomorphisms. Thus we may also use $\text{Hom}_\mathbb{Z}(\Theta, \mathbb{Z})$ for space of homomorphisms from $\Theta$ into $\mathbb{Z}$, as commutative groups or modules over $\mathbb{Z}$. This is a commutative group with respect to pointwise addition of functions into $\mathbb{Z}$, or equivalently a module over $\mathbb{Z}$.

Elements of $\text{Hom}(\Theta, \mathbb{Z})$ are uniquely determined by their restrictions to $\Delta$. Every $\mathbb{Z}$-valued function on $\Delta$ can be extended to a homomorphism from $\Theta$
into \( Z \). Similarly, linear functionals on \( V \) are uniquely determined by their restrictions to \( \Delta \), and every real-valued function on \( \Delta \) can be extended to a linear functional on \( V \). If \( \lambda \) is a linear functional on \( V \) that maps \( \Theta \) into \( Z \), then the restriction of \( \lambda \) to \( \Theta \) defines an element of \( \text{Hom}(\Theta, Z) \). Every element of \( \text{Hom}(\Theta, Z) \) is the restriction to \( \Theta \) of a unique linear functional on \( V \) that maps \( \Theta \) into \( Z \), which can be obtained from its restriction to \( \Delta \), as before.

Thus \( \text{Hom}(\Theta, Z) \) can be identified with the set of linear functionals on \( V \) that map \( \Theta \) into \( Z \). This is a subgroup of the dual space \( V' \), as a commutative group with respect to addition. If \( \alpha \in \Phi \), then \( \lambda_\alpha \) maps \( \Phi \) into \( Z \), by the definition of a root system. This means that \( \lambda_\alpha \) maps \( \Theta \) into \( Z \), so that the restriction of \( \lambda_\alpha \) to \( \Theta \) is an element of \( \text{Hom}(\Theta, Z) \).

Let \( \Theta_{\Phi'} \) be the subgroup of \( V' \), as a commutative group with respect to addition, generated by \( \Phi' \). Note that every element of \( \Theta_{\Phi'} \) maps \( \Theta \) into \( Z \), so that its restriction to \( \Theta \) is an element of \( \text{Hom}(\Theta, Z) \).

As before, \( \Theta_{\Phi'} \) consists of linear combinations of elements of a base for \( \Phi' \) with coefficients in \( Z \). If \( \Phi \) is reduced as a root system in \( V \), then

\[(23.1.3) \quad \Delta' = \{\lambda_\alpha : \alpha \in \Delta\}\]

is a base for \( \Phi' \), as in Section 19.13.

In the types of situations considered in Section 22.11, one can ask for a homomorphism from \( \Theta_{\Phi'} \) into \( A_0 \), as commutative groups with respect to addition, that sends \( \lambda_\alpha \) to \( h_\alpha \) for every \( \alpha \in \Phi \). In particular, this happens in the type of situation considered in Section 22.1, as in Section 22.2.

### 23.2 Bases and submodules

Let us return to the type of situation considered in Section 22.6, and look at some additional properties related to a base for the root system. Thus we let \( V \) be a vector space over the real numbers of positive finite dimension, and let \( \Phi \) be a root system in \( V \). Also let \( \Delta \) be a base for \( \Phi \), and let \( \Phi^+ = \Phi^{\Delta^+} \) be the set of positive roots in \( \Phi \) with respect to \( \Delta \). Let \( \Theta \) be the subgroup of \( V \), as a commutative group with respect to addition, generated by \( \Phi \), as before. Of course, this is the same as the subgroup of \( V \) generated by \( \Delta \).

Let \( k \) be a commutative ring with a multiplicative identity element, and let \((A, [\cdot, \cdot], a)\) be a Lie algebra over \( k \). Suppose that \( A_\alpha \) is a submodule of \( A \), as a module over \( k \), for each \( \alpha \in \Phi \cup \{0\} \), and that \( A \) corresponds to the direct sum of the \( A_\alpha \)’s, as a module over \( k \). Put \( A_\alpha = \{0\} \) when \( \alpha \in \Theta \setminus (\Phi \cup \{0\}) \), as before, and suppose that \([A_\alpha, A_\beta] \subseteq A_{\alpha+\beta} \) for every \( \alpha, \beta \in \Theta \). Of course, this implies that

\[(23.2.1) \quad [A_\alpha, A_{-\alpha}] \subseteq A_0\]

for every \( \alpha \in \Theta \). In particular, \( A_0 \) is a Lie subalgebra of \( A \), as before.

If \( \alpha, \beta \in \Delta \), then \( \alpha - \beta \notin \Phi \), by the definition of a base for a root system. If \( \alpha \neq \beta \) too, then \( \alpha - \beta \notin \Phi \cup \{0\} \), so that

\[(23.2.2) \quad [A_\alpha, A_{-\beta}] = \{0\}\]
23.3. SOME GENERATING SUBMODULES

This corresponds to part of the proposition on p96 of [13], and part of part (b) of Theorem 6 on p48 of [23].

If \( \alpha \in \Phi \), then let \( \sigma_\alpha \) be the symmetry on \( V \) with vector \( \alpha \) that maps \( \Phi \) onto itself, and let \( \lambda_\alpha \) be the corresponding linear functional on \( V \), so that \( \sigma_\alpha(v) = v - \lambda_\alpha(v) \alpha \) for every \( v \in V \). Let \( \alpha, \beta \in \Delta \) be given, with \( \alpha \neq \beta \). Remember that \( \lambda_\alpha(\beta) \in \mathbb{Z} \), by definition of a root system, and that \( \lambda_\alpha(\beta) \leq 0 \) in this case, as in Section 20.2. Observe that

\[
\beta - \lambda_\alpha(\beta) \alpha + \alpha = \sigma_\alpha(\beta - \alpha),
\]

because \( \sigma_\alpha(\alpha) = -\alpha \). This implies that

\[
\beta - \lambda_\alpha(\beta) \alpha + \alpha \in \mathbb{Z} \setminus (\Phi \cup \{0\}),
\]

because \( \lambda_\alpha(\beta) \in \mathbb{Z} \) and \( \alpha - \beta \notin \Phi \cup \{0\} \), as before. It follows that

\[
A_{\beta - \lambda_\alpha(\beta) \alpha + \alpha} = \{0\}
\]

under these conditions. Similarly,

\[
-\beta - \lambda_\alpha(\beta) (-\alpha) - \alpha = -(\beta - \lambda_\alpha(\beta) \alpha + \alpha) \in \mathbb{Z} \setminus (\Phi \cup \{0\}),
\]

so that

\[
A_{-\beta - \lambda_\alpha(\beta) (-\alpha) - \alpha} = \{0\}.
\]

This corresponds to another part of the proposition on p96 of [13], and to part (c) of Theorem 6 on p48 of [23].

Let \( \alpha, \beta \in \Theta \) be given, so that

\[
[[A_\alpha, A_\beta], A_{-\alpha - \beta}] \subseteq [A_{\alpha + \beta}, A_{-\alpha - \beta}] \subseteq A_0.
\]

Using the Jacobi identity, we get that

\[
[[A_\alpha, A_\beta], A_{-\alpha - \beta}] \subseteq [[A_\beta, A_{-\alpha - \beta}], A_\alpha] + [[A_{-\alpha - \beta}, A_\alpha], A_\beta].
\]

This implies that

\[
[[A_\alpha, A_\beta], A_{-\alpha - \beta}] \subseteq [A_\alpha, A_{-\alpha}] + [A_\beta, A_{-\beta}].
\]

23.3 Some generating submodules

Let us continue with the same notation and hypotheses as in the previous section. If \( \Psi_0 \subseteq \Theta \), then let \( A(\Psi_0) \) be the subset of \( A \) consisting of finite sums of elements of \( A_\alpha, \alpha \in \Psi_0 \), as before. This is a submodule of \( A \), as a module over \( k \).

Suppose that for each \( \alpha \in \Delta \) and \( \beta \in \Phi^+ \) with \( \alpha + \beta \in \Phi \) we have that

\[
[A_\alpha, A_\beta] = A_{\alpha + \beta}.
\]
Of course, \( \alpha + \beta \in \Phi^+ \) in this case. This condition holds in the situation considered in Section 22.1, as in Section 17.9. Note that \( \alpha + \beta \neq 0 \) when \( \alpha, \beta \) are elements of \( \Phi^+ \), because of the nonnegativity of the coefficients of \( \alpha \) and \( \beta \) when expressed as linear combinations of elements of \( \Delta \). If \( \alpha + \beta \not\in \Phi \), then \( A_{\alpha+\beta} = \{0\} \), by construction, and (23.3.1) holds automatically, by hypothesis.

Let \( \beta \in \Phi^+ \) be given, and remember that \( \beta \) can be expressed as \( \sum_{j=1}^{r} \alpha_j \) for some \( \alpha_1, \ldots, \alpha_r \in \Delta \), where

\[
(23.3.2) \quad \beta_l = \sum_{j=1}^{l} \alpha_j \in \Phi
\]

for each \( l = 1, \ldots, r \), as in Section 19.12. If \( l < r \), then we get that

\[
(23.3.3) \quad A_{\beta_{l+1}} = [A_{\alpha_{l+1}}, A_{\beta_l}],
\]

as in (23.3.1), because \( \beta_l \in \Phi^+ \).

Remember that \( A(\Phi^+) \) is a Lie subalgebra of \( A \), as in Section 22.6. If (23.3.1) holds for every \( \alpha \in \Delta \) and \( \beta \in \Phi^+ \) with \( \alpha + \beta \in \Phi \), then

\[
(23.3.4) \quad A(\Phi^+) \text{ is generated by the submodules } A_\alpha, \alpha \in \Delta, \text{ as a Lie algebra over } k.
\]

More precisely, if \( \beta \in \Phi^+ \), then one can use (23.3.3) to get that \( A_\beta \) is contained in the Lie subalgebra of \( A \) generated by \( A_\alpha, \alpha \in \Delta \). This corresponds to part of the proof of the proposition on p74 of [13], and is related to part of part (a) of Theorem 6 on p48 of [23].

Similarly, suppose that

\[
(23.3.5) \quad [A_{-\alpha}, A_{-\beta}] = A_{-\alpha-\beta}
\]

for every \( \alpha \in \Delta \) and \( \beta \in \Phi^+ \) with \( \alpha + \beta \in \Phi \). This also holds in the situation considered in Section 22.1, as in Section 17.9. If \( \alpha + \beta \not\in \Phi \), then \( -\alpha - \beta \not\in \Phi \), so that \( A_{-\alpha-\beta} = \{0\} \), and (23.3.5) holds automatically, by hypothesis. If \( \beta \in \Phi^+ \), and \( \alpha_1, \ldots, \alpha_r \in \Delta \), \( \beta_l \in \Phi^+ \) are as in (23.3.2), then we get that

\[
(23.3.6) \quad A_{-\beta_{l+1}} = [A_{-\alpha_{l+1}}, A_{-\beta_l}]
\]

when \( l < r \).

We also have that \( A(-\Phi^+) \) is a Lie subalgebra of \( A \), as in Section 22.6 again. If (23.3.5) holds for every \( \alpha \in \Delta \) and \( \beta \in \Phi^+ \) with \( \alpha + \beta \in \Phi \), then

\[
(23.3.7) \quad A(-\Phi^+) \text{ is generated by the submodules } A_{-\alpha}, \alpha \in \Delta, \text{ as a Lie algebra over } k.
\]

This can be obtained from (23.3.6), as before. This corresponds to another part of the proof of the proposition on p74 of [13], and is related to another part of part (a) of Theorem 6 on p48 of [23].
23.4. SOME GENERATORS

If \( \alpha \in \Delta \), \( \beta \in \Phi^+ \), and (23.3.1) holds, then

\[
[A_{\alpha+\beta}, A_{-\alpha-\beta}] \subseteq [A_{\alpha}, A_{-\alpha}] + [A_\beta, A_{-\beta}],
\]

by (23.2.10). Suppose that (23.3.1) holds for every \( \alpha \in \Delta \) and \( \beta \in \Phi^+ \) with \( \alpha + \beta \in \Phi \) again. If \( \beta \in \Phi^+ \), then

\[
[A_{\beta}, A_{-\beta}] \text{ is contained in the submodule of } A_0
\]

generated by \([A_\alpha, A_{-\alpha}], \alpha \in \Delta\), as a module over \(k\).

To see this, one can express \( \beta \) as in (23.3.2) with \( l = r \), as before. If \( l < r \), then

\[
[A_{\beta_{l+1}}, A_{-\beta_{l+1}}] \subseteq [A_{\alpha_{l+1}}, A_{-\alpha_{l+1}}] + [A_\beta, A_{-\beta}],
\]

by (23.3.3) and (23.3.8). One can use this repeatedly to get (23.3.9). Of course, one can use an analogous argument when (23.3.5) holds for every \( \alpha \in \Delta \) and \( \beta \in \Phi^+ \) with \( \alpha + \beta \in \Phi \), instead of (23.3.1).

In the situation considered in Section 22.1,

\[
A_0 \text{ is generated by } [A_\alpha, A_{-\alpha}], \alpha \in \Phi^+, \text{ as a module over } k.
\]

as in (22.11.7). More precisely,

\[
A_0 \text{ is generated by } [A_\alpha, A_{-\alpha}], \alpha \in \Delta, \text{ as a module over } k.
\]

in this case, as in (22.11.8). If (23.3.9) holds for every \( \beta \in \Phi^+ \), then (23.3.11) implies (23.3.12).

Suppose that (23.3.1) and (23.3.5) hold for every \( \alpha \in \Delta \) and \( \beta \in \Phi^+ \) with \( \alpha + \beta \in \Phi \). If (23.3.11) holds too, then we get that

\[
A \text{ is generated by the submodules}
\]

\[
A_\alpha \text{ and } A_{-\alpha}, \alpha \in \Delta, \text{ as a Lie algebra over } k.
\]

This corresponds to the proposition on p74 of [13], which is related to another part of part (a) of Theorem 6 on p48 of [23].

23.4 Some generators

Let us continue with the same notation and basic hypotheses as in the previous two sections. Suppose now that for each \( \alpha \in \Delta \), there is an \( x_\alpha \in A_\alpha \) such that

\[
A_\alpha \text{ is generated by } x_\alpha \text{ as a module over } k,
\]

so that every element of \( A_\alpha \) can be expressed as a multiple of \( x_\alpha \) by an element of \( k \). This holds in the situation considered in Section 22.1, as in Section 17.7.

If \( \alpha \in \Delta \) and \( \beta \in \Phi^+ \) satisfy \( \alpha + \beta \in \Phi \), then (23.3.1) is the same as saying that

\[
\text{ad}_{x_\alpha}(A_\beta) = A_{\alpha+\beta}.
\]
Suppose that this holds for all such $\alpha$ and $\beta$, let $\beta \in \Phi^+$ be given, and let $\alpha_1, \ldots, \alpha_r \in \Delta$ be as before, so that (23.3.2) holds for every $l = 1, \ldots, r$, and $\beta_r = \beta$. If $l < r$, then
\begin{equation}
A_{\beta_l+1} = \text{ad}_{x_{\alpha_{l+1}}} (A_{\beta_l}),
\end{equation}
by (23.4.2). This implies that $A_{\beta_l}$ is generated by a single element as a module over $k$ for each $l = 1, \ldots, r$. In particular, $A_{\beta}$ is generated by a single element, as a module over $k$, because $\beta = \beta_r$.

In this situation, (23.3.4) is the same as saying that
\begin{equation}
A_{\beta_l+1} = \text{ad}_{x_{\alpha_{l+1}}} (A_{\beta_l}),
\end{equation}

(23.4.3)

This corresponds to part of part (a) of Theorem 6 on p48 of [23].

Similarly, suppose that for each $\alpha_2 \in \Delta$ there is a $y_\alpha \in A_{-\alpha}$ such that
\begin{equation}
A_{-\alpha} \text{ is generated by } y_\alpha \text{ as a module over } k.
\end{equation}

(23.4.5)

This holds in the situation considered in Section 22.1, as in Section 17.7 again. If $\alpha \in \Delta$, $\beta \in \Phi^+$, and $\alpha + \beta \in \Phi$, then (23.3.5) is the same as saying that
\begin{equation}
\text{ad}_{y_\alpha} (A_{-\beta}) = A_{-\alpha - \beta}.
\end{equation}

(23.4.6)

If this holds for all such $\alpha$ and $\beta$, then we get that $A_{-\beta}$ is generated by a single element as a module over $k$ for every $\beta \in \Phi^+$, as before.

In this case, (23.3.7) is the same as saying that
\begin{equation}
A(-\Phi^+) \text{ is generated by } y_\alpha, \alpha \in \Delta, \text{ as a Lie algebra over } k.
\end{equation}

(23.4.7)

This corresponds to another part of part (a) of Theorem 6 on p48 of [23].

Suppose that (23.4.1) and (23.4.5) hold for every $\alpha \in \Delta$, so that
\begin{equation}
[A_{-\alpha}, A_{-\beta}] \text{ is generated by } [x_\alpha, y_\beta]_A \text{ as a module over } k
\end{equation}

for every $\alpha \in \Delta$. In this situation, (23.3.12) is the same as saying that
\begin{equation}
A_0 \text{ is generated by } [x_\alpha, y_\beta]_A, \alpha \in \Delta, \text{ as a module over } k.
\end{equation}

(23.4.8)

If (23.4.2) and (23.4.6) hold for every $\alpha \in \Delta$ and $\beta \in \Phi^+$ with $\alpha + \beta \in \Phi$, and if (23.4.9) holds, then we get that
\begin{equation}
A \text{ is generated by } x_\alpha \text{ and } y_\alpha, \alpha \in \Delta, \text{ as a Lie algebra over } k.
\end{equation}

(23.4.10)

This corresponds to the second formulation of the proposition on p74 of [13], and to another part of part (a) of Theorem 6 on p48 of [23].

23.5 Some relations

Let us continue with the same notation and basic hypotheses as in the previous three sections. If $\alpha \in \Phi$, then let $\sigma_\alpha$ be the symmetry on $V$ with vector $\alpha$ that
maps $\Phi$ onto itself, and let $\lambda_\alpha$ be the corresponding linear functional on $V$. Thus $\lambda_\alpha$ is equal to 0 on the hyperplane fixed by $\sigma_\alpha$ and satisfies $\lambda_\alpha(\alpha) = 2$, so that $\sigma_\alpha(v) = v - \lambda_\alpha(v) \alpha$ for every $v \in V$. Suppose that $\phi$ is a group homomorphism from $\Theta$ into $\text{Hom}_k(A_0, k)$, as a commutative group with respect to addition, such that $\text{ad}_w(x) = \phi_\alpha(x)$ for every $w \in A_0$ and $x \in A_\alpha$, $\alpha \in \Phi \cup \{0\}$, where $\phi_\alpha = \phi(\alpha)$, as before. In particular, this means that $A_0$ is commutative as a Lie subalgebra of $A$.

Suppose that for each $\alpha \in \Delta$, $x_\alpha \in A_\alpha$ and $y_\alpha \in A_{-\alpha}$, and put
\begin{equation}
(23.5.1) \quad h_\alpha = [x_\alpha, y_\alpha]_A
\end{equation}
which is an element of $A_0$. We also suppose that
\begin{equation}
(23.5.2) \quad \phi_\beta(h_\alpha) = \lambda_\alpha(\beta) \cdot 1
\end{equation}
in $k$ for every $\alpha \in \Delta$ and $\beta \in \Phi$. Remember that $\lambda_\alpha(\beta) \in \mathbb{Z}$ for every $\alpha, \beta \in \Phi$.

In the situation considered in Section 22.1, we can choose $x_\alpha$, $y_\alpha$ so that (23.5.1) satisfies (23.5.2), as in Section 22.11. These choices of $x_\alpha$ and $y_\alpha$ satisfy the other conditions mentioned in the previous section in that situation too, as before.

If $\alpha, \beta \in \Delta$, then
\begin{equation}
(23.5.3) \quad [h_\alpha, h_\beta]_A = 0,
\end{equation}
because $A_0$ is commutative as a Lie algebra over $k$. If $\alpha \neq \beta$, then
\begin{equation}
(23.5.4) \quad [x_\alpha, y_\beta]_A = 0,
\end{equation}
as in (23.2.2). We also have that
\begin{equation}
(23.5.5) \quad [h_\alpha, x_\beta]_A = \phi_\beta(h_\alpha) x_\beta = \lambda_\alpha(\beta) \cdot x_\beta
\end{equation}
for every $\alpha, \beta \in \Delta$, using (23.5.2) and the fact that $x_\beta \in A_\beta$, by construction. Similarly,
\begin{equation}
(23.5.6) \quad [h_\alpha, y_\beta]_A = \phi_{-\beta}(h_\alpha) y_\beta = -\lambda_\alpha(\beta) \cdot y_\beta
\end{equation}
for every $\alpha, \beta \in \Delta$, because $y_\beta \in A_{-\beta}$. These relations, including (23.5.1), are called the Weyl relations, as in part (b) of Theorem 6 on p48 of [23].

Let $\alpha, \beta \in \Delta$ be given, with $\alpha \neq \beta$. Remember that $\lambda_\alpha(\beta) \leq 0$, as in Section 20.2. Observe that
\begin{equation}
(23.5.7) \quad \text{ad}^{-\lambda_\alpha(\beta)+1}_{x_\alpha}(x_\beta) = 0,
\end{equation}
by (23.2.5). Similarly,
\begin{equation}
(23.5.8) \quad \text{ad}^{-\lambda_\alpha(\beta)+1}_{y_\alpha}(y_\beta) = 0,
\end{equation}
by (23.2.7). This corresponds to part (c) of Theorem 6 on p48 of [23], and the Weyl relations together with these additional relations are in the proposition on p96 of [13].
23.6 Isomorphisms and diagonalizability

Let \( k \) be a field, and let \((A_1, [\cdot, \cdot]_{A_1})\) and \((A_2, [\cdot, \cdot]_{A_2})\) be finite-dimensional Lie algebras over \( k \). Also let \( T \) be a Lie algebra isomorphism from \( A_1 \) onto \( A_2 \). Suppose that \( B_1 \) is a Lie subalgebra of \( A_1 \) such that every element of \( B_1 \) is ad-diagonalizable as an element of \( A_1 \). This implies that \( B_1 \) is commutative as a Lie algebra over \( k \), as in Section 17.1. It follows that \( B_2 = T(B_1) \) has the analogous properties in \( A_2 \).

Let \( B_1', B_2' \) be the duals of \( B_1, B_2 \), respectively, as vector spaces over \( k \). The restriction \( T_B \) of \( T \) to \( B_1 \) is a one-to-one linear mapping from \( B_1 \) onto \( B_2 \). If \( \alpha_2 \in B_2' \), then

\[
T_B'(\alpha_2) = \alpha_2 \circ T_B \in B_1',
\]
as usual. This defines the dual linear mapping \( T_B' \) from \( B_2' \) into \( B_1' \). More precisely, \( T_B' \) is a one-to-one mapping from \( B_2' \) onto \( B_1' \) in this situation.

If \( \alpha_j \in B_j', j = 1, 2 \), then let \( A_{j, \alpha_j} \) be the set of \( x_j \in A_j \) such that

\[
\text{ad}_{A_j, w_j}(x_j) = [w_j, x_j]_{A_j} = \alpha_j(w_j) x_j
\]
for every \( w_j \in B_j \), as before. Thus \( A_{j, \alpha_j} \) is a linear subspace of \( A_j \), as a vector space over \( k \). Observe that \( x_1 \in A_1, w_1 \in B_1 \), and \( \alpha_1 \in B_1' \) satisfy

\[
[w_1, x_1]_{A_1} = \alpha_1(w_1) x_1
\]
exactly when

\[
[T(w_1), T(x_1)]_{A_2} = \alpha_1(w_1) T(x_1).
\]

Let \( \alpha_2 \in B_2' \) be given, and let us apply this to \( \alpha_1 = T_B'(\alpha_2) \). This implies that \( x_1 \in A_1 \) and \( w_1 \in B_1 \) satisfy

\[
[w_1, x_1]_{A_2} = (T_B'(\alpha_2))(w_1) x_1
\]
exactly when

\[
[T(w_1), T(x_1)]_{A_2} = \alpha_2(T(w_1)) T(x_1),
\]
because \( (T_B'(\alpha_2))(w_1) = \alpha_2(T_B(w_1)) = \alpha_2(T(w_1)) \).

Using this, it is easy to see that

\[
T(A_1, T_B'(\alpha_2)) = A_2, \alpha_2
\]
for every \( \alpha_2 \in B_2' \). Put

\[
\Phi_{j, B_j} = \{ \alpha_j \in B_j' : \alpha_j \neq 0 \text{ and } A_{j, \alpha_j} \neq \{0\} \}
\]
for \( j = 1, 2 \), as before. It follows that

\[
T_B'(\Phi_{2, B_2}) = \Phi_{1, B_1},
\]
by (23.6.8). Remember that \( \Phi_{j, B_j} \) has finitely many elements, as in Section 17.2.

Suppose now that \( k \) has characteristic 0. Let \( E_j, Q \) be the linear subspace of \( B_j^\prime \), as a vector space over \( Q \), spanned by \( \Phi_{j, B_j} \) for \( j = 1, 2 \), as before. Using (23.6.10), we get that

\[
T_B^\prime(E_2, Q) = E_1, Q.
\]

Thus the restriction of \( T_B^\prime \) to \( E_2, Q \) is a one-to-one mapping onto \( E_1, Q \) that is linear over \( Q \).

Let \( E_j, R \) be the vector space over \( R \) obtained from \( E_j, Q \) as in Section 17.13, for \( j = 1, 2 \). More precisely, \( E_j, Q \) corresponds to a linear subspace of \( E_j, R \), as a vector space over \( R \), and any basis for \( E_j, Q \), as a vector space over \( Q \), is also a basis for \( E_j, R \), as a vector space over \( R \). The restriction of \( T_B^\prime \) to \( E_2, Q \) leads to a one-to-one mapping from \( E_2, R \) onto \( E_1, R \) that is linear over \( R \). This mapping agrees with \( T_B^\prime \) on \( E_2, Q \), and in particular maps \( 2, B_2 \) onto \( 1, B_1 \).

If \( 1, B_1 \) is a root system in \( E_1, R \), then \( 2, B_2 \) is a root system in \( E_2, R \). In this case, the linear mapping from \( E_2, R \) onto \( E_1, R \) obtained from \( T_B^\prime \) as in the preceding paragraph is an isomorphism between these root systems. Remember that \( \Phi_{j, B_j} \) is a root system in \( E_j, R \) when \( A_j, B_j \) satisfy the conditions in Section 22.1.

### 23.7 Automorphisms and \( SL_2(k) \)

Let \( k \) be a field of characteristic 0, and let \( (A, [\cdot, \cdot], A) \) be a finite-dimensional Lie algebra over \( k \). If \( w \in A \) is ad-nilpotent, then the exponential of \( \text{ad}_w \) defines a Lie algebra automorphism of \( A \), as in Section 14.11. If \( u \in A \) and \( [w, u], A = 0 \), then it is easy to see that

\[
(\exp \text{ad}_w)(u) = u.
\]

Let \( x, y \) be elements of \( A \), and put

\[
h = [x, y], A.
\]

Suppose that

\[
[h, x], A = 2 \cdot x, \quad [h, y], A = -2 \cdot y.
\]

Let \( C \) be the linear span of \( x, y, h \) in \( A \), which is a Lie subalgebra of \( A \) that is isomorphic to \( SL_2(k) \), as a Lie algebra over \( k \). Using (23.7.3), we get that \( x \) and \( y \) are ad-nilpotent as elements of \( A \), as in Section 14.2. More precisely, this uses the adjoint representation of \( A \).

Of course, \( -y \) is ad-nilpotent as an element of \( A \) too. Thus the exponentials of \( \text{ad}_x \) and \( -\text{ad}_y \) define Lie algebra automorphisms on \( A \), as before. It follows that

\[
\theta = (\exp \text{ad}_x) \circ (\exp -\text{ad}_y) \circ (\exp \text{ad}_x)
\]

is a Lie algebra automorphism on \( A \) too. If \( u \in A \) satisfies

\[
[x, u], A = [y, u], A = 0,
\]
then
\[(\exp \text{ad}_x)(u) = (\exp -\text{ad}_y)(u) = u,\]
as in (23.7.1). This implies that
\[(\exp \text{ad}_x)(u) = (\exp \text{ad}_y)(u) = u;\]  
(23.7.6)

Clearly \(\text{ad}_x\) and \(\text{ad}_y\) map \(C\) into itself, because \(C\) is a Lie subalgebra of \(A\) that contains \(x\) and \(y\). The restrictions of \(\text{ad}_x = \text{ad}_{A,x}\) and \(\text{ad}_y = \text{ad}_{A,y}\) to \(C\) are the same as \(\text{ad}_{C,x}\) and \(\text{ad}_{C,y}\), respectively, corresponding to the adjoint representation on \(C\). It is easy to see that the restrictions of \(\exp \text{ad}_x = \exp \text{ad}_{A,x}\) and \(\exp(-\text{ad}_y) = \exp(-\text{ad}_{A,y})\) to \(C\) are the same as \(\exp \text{ad}_{C,x}\) and \(\exp(-\text{ad}_{C,y})\). This means that the restriction of \(\theta\) to \(C\) is the same as
\[(\exp \text{ad}_{C,x}) \circ (\exp \text{ad}_{C,y}) \circ (\exp \text{ad}_{C,x});\]  
(23.7.8)

Using this, we get that
\[\theta(x) = -y, \quad \theta(y) = -x, \quad \theta(h) = -h.\]  
(23.7.9)

More precisely, we can take \(\theta\) to be as in (23.7.8) here, as in the preceding paragraph. This permits one to obtain (23.7.9) as in Section 15.7, because \(C\) is isomorphic to \(\mathfrak{sl}_2(k)\) as a Lie algebra over \(k\), with its usual basis. Note that (23.7.4) is an inner automorphism of \(A\), as in Section 14.11.

### 23.8 Automorphisms and roots

Let \(k\) be a field of characteristic 0 again, and let \((A, [\cdot, \cdot]_A)\) be a finite-dimensional Lie algebra over \(k\). Also let \(B\) be a Lie subalgebra of \(A\), all of whose elements are ad-diagonalizable as elements of \(A\). If \(\alpha\) is an element of the dual \(B'\) of \(B\), as a vector space over \(k\), then we let \(A_\alpha\) be the set of \(x \in A\) such that \([w, x]_A = \alpha(w)x\) for every \(w \in B\), as before. The set of \(\alpha \in B'\) such that \(\alpha \neq 0\) and \(A_\alpha \neq \{0\}\) is denoted \(\Phi_B\), as usual.

Let \(\alpha \in \Phi_B\) be given, and suppose that \(x_\alpha \in A_\alpha\) and \(y_\alpha \in A_{-\alpha}\) have the property that \(h_\alpha = [x_\alpha, y_\alpha]_A\) satisfies \(\alpha(h_\alpha) = 2\). Thus \([h_\alpha, x_\alpha]_A = \alpha(h_\alpha)x_\alpha = 2x_\alpha\), and \([h_\alpha, y_\alpha]_A = -\alpha(h_\alpha)y_\alpha = -2y_\alpha\). In particular, \(x_\alpha\) and \(y_\alpha\) are ad-nilpotent as elements of \(A\), as in Sections 14.2 and 17.2. This implies that \(-y_\alpha\) is ad-nilpotent as an element of \(A\), and that the exponentials of \(\text{ad}_{x_\alpha}\) and \(-\text{ad}_{y_\alpha}\) define Lie algebra automorphisms of \(A\), as in Section 14.11. Note that there are such \(x_\alpha, y_\alpha\) in the situation considered in Section 22.1, as in Section 17.6.

Put
\[(23.8.1) \quad \theta_\alpha = (\exp \text{ad}_{x_\alpha}) \circ (\exp -\text{ad}_{y_\alpha}) \circ (\exp \text{ad}_{x_\alpha}),\]
which defines a Lie algebra automorphism on \(A\). If \(u \in A\) satisfies \([x_\alpha, u]_A = [y_\alpha, u]_A = 0\), then
\[(23.8.2) \quad \theta_\alpha(u) = u,\]
23.8. AUTOMORPHISMS AND ROOTS

as in the previous section. In particular, this holds for every $u \in B$ with $\alpha(u) = 0$, because $x_\alpha \in A_\alpha$ and $y_\alpha \in A_{-\alpha}$. We also have that
\[(23.8.3)\quad \theta_\alpha(h_\alpha) = -h_\alpha,\]
as before. This implies that
\[(23.8.4)\quad \theta_\alpha(B) = B,\]
because $B$ is spanned by the kernel of $\alpha$ and $h_\alpha$.

Let $\theta_{\alpha,B}$ be the restriction of $\theta_\alpha$ to $B$, which is a one-to-one linear mapping from $B$ onto itself. If $\beta \in B'$, then
\[(23.8.5)\quad \theta'_{\alpha,B}(\beta) = \beta \circ \theta_{\alpha,B} \in B',\]
and $\theta'_{\alpha,B}$ is a one-to-one linear mapping from $B'$ onto itself. As in Section 23.6,
\[(23.8.6)\quad \theta_\alpha(A_{\theta_{\alpha,B}(\beta)}) = A_\beta\]
for every $\beta \in B'$. This implies that
\[(23.8.7)\quad \theta'_{\alpha,B}(\Phi_B) = \Phi_B,\]
as before.

It is easy to see that
\[(23.8.8)\quad \theta_\alpha(u) = u - \alpha(u) h_\alpha\]
for every $u \in B$, using (23.8.2), (23.8.3), and the fact that $\alpha(h_\alpha) = 2$. If $\beta \in B'$, then we get that
\[(23.8.9)\quad \theta'_{\alpha,B}(\beta) = \beta - \beta(h_\alpha) \alpha.\]

Let $E_Q$ be the linear subspace of $B'$, as a vector space over $Q$, spanned by $\Phi_B$, as before. Thus
\[(23.8.10)\quad \theta'_{\alpha,B}(E_Q) = E_Q,\]
as in Section 23.6, by (23.8.7).

Suppose that for each $\beta \in \Phi_B$, $\beta(h_\alpha)$ corresponds to an element of $Q$ with respect to the standard embedding into $k$, which holds in the situation considered in Section 22.1. This implies that $\beta(h_\alpha)$ is in the image of $Q$ in $k$ when $\beta \in E_Q$, by definition of $E_Q$. Let $\lambda_\alpha$ be the linear functional on $E_Q$, as a vector space over $Q$, corresponding to $\beta \mapsto \beta(h_\alpha)$. If $\beta \in E_Q$, then (23.8.9) can be expressed as
\[(23.8.11)\quad \theta'_{\alpha,B}(\beta) = \beta - \lambda_\alpha(\beta) \alpha.\]

Let $E_R$ be the vector space over $R$ obtained from $E_Q$ as in Sections 17.13 and 23.6, so that $E_Q$ corresponds to a linear subspace of $E_R$, as a vector space over $Q$, and any basis for $E_Q$, as a vector space over $Q$, is a basis for $E_R$ too, as a vector space over $R$. The restriction of $\theta'_{\alpha,B}$ to $E_Q$ leads to a one-to-one mapping from $E_R$ onto itself that is linear over $R$, as in Section 23.6.

There is a natural extension of $\lambda_\alpha$ to a linear functional on $E_R$, that we shall denote as $\lambda_\alpha$ as well. The mapping on $E_R$ corresponding to the restriction
of $\theta'_{\alpha_B}$ to $E_Q$ can be given by the same expression as in (23.8.11), using this extension of $\lambda_{\alpha}$ to $E_R$. This defines a symmetry on $E_R$ with vector $\alpha$ that maps $\Phi_B$ onto itself.

In the situation considered in Section 22.1, $\Phi_B$ is a root system in $E_R$, and every $\alpha \in \Phi_B$ can be handled in this way. It follows that every element of the Weyl group of $\Phi_B$ corresponds to a Lie algebra automorphism of $A$ that maps $B$ onto itself as in Section 23.6 in this case. This corresponds to the first part of the remark on p47 of [23]. This also corresponds to some remarks on pp77 of [13]. More precisely, every element of the Weyl group of $\Phi_B$ corresponds to an inner automorphism of $A$ that maps $B$ into itself in this way, because (23.8.1) is an inner automorphism of $A$, as in the previous section.

23.9 Automorphisms and bases for $\Phi_B$

We would like to consider some conditions combining those in Section 22.1 and the previous section, and which hold in the situation described in Section 22.1, as usual. We begin as in the previous section, so that $(A, [\cdot, \cdot]_A)$ is a finite-dimensional Lie algebra over a field $k$ of characteristic 0. Let $B$, $A_\alpha$ for $\alpha \in B'$, and $\Phi_B$ be as before. Remember that $B$ is commutative as a Lie subalgebra of $A$, as in Section 17.1, and that $A$ corresponds to the direct sum of the linear subspaces $A_\alpha$, $\alpha \in \Phi_B \cup \{0\}$, as a vector space over $k$, as in Section 17.2. We also have that $[A_\alpha, A_\beta] \subseteq A_{\alpha+\beta}$ for every $\alpha, \beta \in B'$, as before.

Of course, $B \subseteq A_0$, because $B$ is commutative as a Lie subalgebra of $A$. We ask that $B = A_0$, which will be compatible with Section 22.12. This is the same as saying that $B$ is equal to its centralizer in $A$, as usual.

Suppose that for each $\alpha \in \Phi_B$, we have elements $x_\alpha$ of $A_\alpha$ and $y_\alpha$ of $A_{-\alpha}$ such that $h_{\alpha} = [x_\alpha, y_\alpha]_A$ satisfies $\alpha(h_{\alpha}) = 2$, as in the previous section. Suppose also that $A_{\alpha}$, $A_{-\alpha}$ are spanned by $x_\alpha$, $y_\alpha$, respectively, as in Section 22.12.

Let $E_Q$ be the linear subspace of $B'$, as a vector space over $Q$, spanned by $\Phi_B$, and let $E_R$ be the vector space over $R$ obtained from $E_Q$ in the usual way, as in the previous section. Suppose that for each $\alpha, \beta \in \Phi_B$, $\beta(h_{\alpha})$ corresponds to an element of $Q$ with respect to the standard embedding into $k$, as in the previous section. If $\alpha \in \Phi_B$, then we get a linear functional $\lambda_{\alpha}$ on $E_Q$, as a vector space over $Q$, corresponding to $\beta \mapsto \beta(h_{\alpha})$, as before. This has a natural extension to a linear functional on $E_R$, which is denoted $\lambda_{\alpha}$, too, as before. This can be used to get a symmetry on $E_R$ with vector $\alpha$ that maps $\Phi_B$ onto itself, as in the previous section.

We would like $\Phi_B$ to be a root system in $E_R$. This amounts to asking that for every $\alpha, \beta \in \Phi_B$, $\beta(h_{\alpha})$ corresponds to an integer with respect to the standard embedding into $k$, so that $\lambda_{\alpha}(\beta) \in Z$. Under these conditions, we can take $V = E_R$ and $\Phi = \Phi_B$ in Section 22.12.

If $\Theta$ is the subgroup of $E_R$, as a commutative group with respect to addition, generated by $\Phi_B$, then $\Theta$ is in fact a subgroup of $E_Q$, and thus of $B'$. Equivalently, $\Theta$ is a subgroup of the dual $A'_0$ of $A_0$, as a vector space over $k$. The group homomorphism $\phi$ from $\Theta$ into $A'_0$ mentioned in Section 22.12 may
23.9. AUTOMORPHISMS AND BASES FOR $\Phi_B$

be taken to be the obvious inclusion mapping of $\Theta$ into $A_0$ here.

If $\Psi_0 \subseteq \Phi_B \cup \{0\}$, then we take $A(\Psi_0)$ to be the linear subspace of $A$ spanned by $A_\alpha$, $\alpha \in \Psi_0$, as in Section 22.12. Let $\Delta$ be a base for $\Phi_B$, as a root system in $E_R$. This leads to the set $\Phi_B^{\Delta,+}$ of positive roots in $\Phi_B$ with respect to $\Delta$, which can be expressed as linear combinations of elements of $\Delta$ with nonnegative coefficients. Remember that $A(\Phi_B^{\Delta,+} \cup \{0\})$ is a solvable Lie subalgebra of $A$, and in fact a Borel subalgebra of $A$, as in Section 22.12. This may be called the standard Borel subalgebra of $A$ associated to $B$ and $\Delta$, as before.

If $\sigma$ is an element of the Weyl group of $\Phi_B$, then there is an inner automorphism $\theta$ of $A$ that corresponds to $\sigma$ as in the previous section. More precisely, $\theta(\hat{B}) = B$, and we let $\theta_B$ be the restriction of $\theta$ to $\hat{B}$, which is a one-to-one linear mapping from $\hat{B}$ onto itself. If $\beta \in B'$, then $\theta_B'(\beta) = \beta \circ \theta_B \in B'$, and $\theta_B'$ is a one-to-one linear mapping from $B'$ onto itself. As before, $\theta$ maps $A_{\theta_B'(\beta)}$ onto $A_\beta$ for every $\beta \in B'$, which implies that $\theta_B'$ maps $\Phi_B$ onto itself. It follows that $\theta_B'$ maps $E_Q$ onto itself, as in the previous section.

The restriction of $\theta_B'$ to $E_R$ leads to a one-to-one linear mapping from $E_R$ onto itself, as before. We can choose $\theta$ so that this linear mapping on $E_R$ is $\sigma$. Indeed, if $\alpha \in \Phi_B$, then the automorphism $\theta_\alpha$ in (23.8.1) corresponds in this way to the symmetry on $E_R$ with vector $\alpha$ that maps $\Phi_B$ onto itself. If $\sigma$ is any element of the Weyl group of $\Phi$, then $\sigma$ can be expressed as the composition of finitely many such symmetries on $E_R$. One can take $\theta$ to be the composition of the corresponding $\theta_\alpha$’s in the other order, because the dual of a composition of linear mappings is the composition of the dual linear mappings in the other order.

If $\Psi_0 \subseteq \Phi_B \cup \{0\}$, then $\theta_B'(\Psi_0) = \sigma(\Psi_0) \subseteq \Phi_B \cup \{0\}$, and

$$\theta(A(\sigma(\Psi_0))) = \theta(A(\theta_B'(\Psi_0))) = A(\Psi_0).$$

Let $\Delta$ be a base for $\Phi_B$ again, so that $\sigma(\Delta)$ is a base for $\Phi_B$ too. Thus $\Phi_B^{\sigma(\Delta),+}$ can be defined in the same way as before, and in fact

$$\sigma(\Phi_B^{\Delta,+}) = \Phi_B^{\sigma(\Delta),+}.\quad (23.9.2)$$

This implies that

$$\theta(A(\Phi_B^{\sigma(\Delta),+})) = A(\Phi_B^{\Delta,+})$$

by (23.9.1).

Suppose that $\Phi_B$ is reduced as a root system in $E_R$. This implies that every base for $\Phi_B$ can be expressed as $\sigma(\Delta)$ for some $\sigma$ in the Weyl group of $\Phi_B$, as in Section 19.14. It follows that the standard Borel subalgebra of $A$ associated to any base of $\Phi_B$ is related to $A(\Phi_B^{\Delta,+} \cup \{0\})$ by an inner automorphism of $A$, as in (23.9.3). This basically corresponds to the second part of Lemma C on p.84 of [13].

More precisely, that result also states that $\theta$ can be taken to be an element of the subgroup $E(A)$ of the group $\text{Int}(A)$ of all inner automorphisms of $A$ defined in Section 24.3. To see this, it is enough to check that if $\alpha \in \Phi_B$, then the automorphism $\theta_\alpha$ in (23.8.1) is an element of $E(A)$. To get this, it is enough
to verify that $\exp \operatorname{ad}_{x_{\alpha}}$ and $\exp(-\operatorname{ad}_{y_{\alpha}})$ are elements of $\mathcal{E}(A)$. This reduces to showing that $x_{\alpha}$ and $y_{\alpha}$ are strongly ad-nilpotent in $A$, in the sense described in Section 24.3, by definition of $\mathcal{E}(A)$. This is easy to do, using the properties of $x_{\alpha}$ and $y_{\alpha}$ mentioned in the previous section.

### 23.10 Sums and diagonalizability

Let $k$ be a field, and let $(A, [\cdot, \cdot]_A)$ be a finite-dimensional Lie algebra over $k$. Also let $B$ be a Lie subalgebra of $A$ such that every element of $B$ is ad-diagonalizable as an element of $A$, which implies that $B$ is commutative as a Lie algebra, as in Section 17.1. As usual, $B'$ denotes the dual of $B$, as a vector space over $k$, and if $\alpha \in B'$, then $A_{\alpha}$ is the set of $x \in A$ such that $[w, x]_A = \alpha(w)x$ for every $w \in B$. Let $\Phi_B$ be the set of $\alpha \in B'$ such that $\alpha \neq 0$ and $A_{\alpha} \neq \{0\}$, as before. Remember that $\Phi_B$ has only finitely many elements, and that $A$ corresponds to the direct sum of the linear subspaces $A_{\alpha}$, $\alpha \in \Phi_B \cup \{0\}$, as a vector space over $k$, as in Section 17.2.

If $C$ is an ideal in $A$, then $\operatorname{ad}_w(C) \subseteq C$ for every $w \in B$, and the restrictions of $\operatorname{ad}_w$, $w \in B$, to $C$ are simultaneously diagonalizable, as in Section 22.9. This means that $C$ corresponds to the direct sum of the linear subspaces $A_{\alpha} \cap C$, as a vector space over $k$, where $\alpha \in \Phi_B \cup \{0\}$ and $A_{\alpha} \cap C \neq \{0\}$, as before.

Let $C_1, \ldots, C_n$ be finitely many nontrivial ideals in $A$, and suppose that $A$ corresponds to their direct sum, as a vector space over $k$, and thus as a Lie algebra. Each $C_j$ corresponds to a direct sum of subspaces of the form $A_{\alpha} \cap C_j$, as a vector space over $k$, in the preceding paragraph. It follows that for each $\alpha \in \Phi_B \cup \{0\}$, $A_{\alpha}$ corresponds to the direct sum of linear subspaces of the form $A_{\alpha} \cap C_j$, $j = 1, \ldots, n$, as a vector space over $k$.

Let us suppose from now on in this section that $B$ is equal to its centralizer in $A$, which means that $B = A_0$. This implies that $B$ corresponds to the direct sum of the linear subspaces

\begin{equation}
B_j = B \cap C_j = A_0 \cap C_j,
\end{equation}

for every $j = 1, \ldots, n$.

If $j \in \{1, \ldots, n\}$, then put

\begin{equation}
\tilde{B}^j = \{\alpha \in B' : \alpha(w_l) = 0 \text{ for every } w_l \in B_j, 1 \leq l \neq j \leq n\}.
\end{equation}

This is a linear subspace of $B'$, and $B'$ corresponds to the direct sum of $\tilde{B}^j$, $1 \leq j \leq n$, as a vector space over $k$. The mapping from $\alpha \in \tilde{B}^j$ to the restriction of $\alpha$ to $B_j$ is a one-to-one linear mapping from $\tilde{B}^j$ onto the dual $B'_j$ of $B_j$ for each
23.10. SUMS AND DIAGONALIZABILITY

Let \( j = 1, \ldots, n \), because \( B \) corresponds to the direct sum of the \( B_j \)'s, \( 1 \leq j \leq n \), as a vector space over \( k \).

Let \( \alpha \in B' \) and \( j \in \{1, \ldots, n\} \) be given, and suppose that

\[
(23.10.4) \quad \text{there is a } w_j \in B_j \text{ such that } \alpha(w_j) \neq 0.
\]

If \( x \in A \), then \( [w_j, x]_\alpha \in C_j \), because \( B_j \subseteq C_j \), and \( C_j \) is an ideal in \( A \). If \( x \in A_\alpha \), then we get that \( \alpha(w_j) x \in C_j \), so that \( x \in C_j \). Thus

\[
(23.10.5) \quad A_\alpha \subseteq C_j
\]

when \( (23.10.4) \) holds.

Suppose that there are also an \( l \in \{1, \ldots, n\} \) and a \( w_l \in B_l \) such that \( j \neq l \) and \( \alpha(w_l) \neq 0 \). This implies that \( A_\alpha \subseteq C_l \), as before, so that \( A_\alpha = \{0\} \), because \( C_j \cap C_l = \{0\} \), by hypothesis.

If \( \alpha \in \Phi_B \), then \( (23.10.4) \) holds for some \( j \in \{1, \ldots, n\} \), because \( \alpha \neq 0 \) and \( B \) corresponds to the direct sum of the \( B_j \)'s. If \( 1 \leq l \leq n \) and \( j \neq l \), then it follows from the remarks in the preceding paragraph that \( \alpha(w_l) = 0 \) for every \( w_l \in B_l \), because \( A_\alpha \neq \{0\} \). This means that \( \alpha \in \tilde{B}_j' \), so that

\[
(23.10.6) \quad \Phi_B \subseteq \bigcup_{j=1}^n \tilde{B}_j'.
\]

Let \( j \in \{1, \ldots, n\} \) be given, and put

\[
(23.10.7) \quad \Phi_{B,j} = \Phi_B \cap \tilde{B}_j'.
\]

If \( \alpha \in \Phi_{B,j} \), then \( (23.10.4) \) holds, because \( \alpha \neq 0 \) and \( \alpha \in \tilde{B}_j' \). This implies that \( (23.10.5) \) holds, as before. Of course,

\[
(23.10.8) \quad \Phi_B = \bigcup_{j=1}^n \Phi_{B,j},
\]

by \( (23.10.6) \). Note that

\[
(23.10.9) \quad \Phi_{B,j} \cap \Phi_{B,l} = \emptyset
\]

when \( 1 \leq j \neq l \leq n \), because \( \tilde{B}_j' \cap \tilde{B}_l' = \{0\} \), by construction.

Let \( \alpha \in \tilde{B}_j' \) be given for some \( j \in \{1, \ldots, n\} \), and suppose that \( \alpha \neq 0 \), so that \( (23.10.4) \) holds. Under these conditions,

\[
(23.10.10) \quad A_\alpha = \{x_j \in C_j : [u_j, x_j]_\alpha = \alpha(u_j) x_j \text{ for every } u_j \in B_j\}.
\]

More precisely, \( A_\alpha \) is contained in the right side, by \( (23.10.5) \). If \( x_j \in C_j \) and \( w_l \in B_l \) for some \( 1 \leq l \neq j \leq n \), then \( [w_l, x_j]_\alpha = 0 \) and \( \alpha(w_l) = 0 \). This implies that the right side of \( (23.10.10) \) is contained in \( A_\alpha \), as desired.

Suppose now that \( k \) has characteristic 0, and that \( \Phi_B \neq \emptyset \). Let \( E_Q \) be the linear subspace of \( B' \), as a vector space over \( Q \), spanned by \( \Phi_B \), as usual.
Similarly, if \( \Phi_{B,j} \neq \emptyset \) for some \( j \in \{1, \ldots, n\} \), then let \( E_{Q,j} \) be the linear subspace of \( B'_j \), as a vector space over \( Q \), spanned by \( \Phi_{B,j} \). Thus

\[
(23.10.11) \quad E_{Q,j} \subseteq E_Q,
\]

and it is easy to see that \( E_Q \) corresponds to the direct sum of the \( E_{Q,j} \)'s for these \( j \)’s, as a vector space over \( Q \).

Let \( E_R \) be the vector space over \( R \) obtained from \( E_Q \) in the usual way, so that \( E_Q \) corresponds to a linear subspace of \( E_R \), as a vector space over \( Q \), and any basis for \( E_Q \), as a vector space over \( Q \), is a basis for \( E_R \), as a vector space over \( R \). Similarly, we can get a vector space \( E_{R,j} \) over \( R \) from \( E_{Q,j} \) for each \( j \in \{1, \ldots, n\} \) such that \( \Phi_{B,j} \neq \emptyset \). More precisely, \( E_{R,j} \) is a linear subspace of \( E_R \), and \( E_R \) corresponds to the direct sum of the \( E_{R,j} \)'s for these \( j \)’s.

Suppose that \( B \) is a root system in \( E_R \). This implies that \( B'_j \) is a root system in \( E_{R,j} \) when \( B'_j \neq \emptyset \). If this happens for more that one \( j \in \{1, \ldots, n\} \), then it follows that \( B \) is reducible as a root system in \( E_R \). This is related to the corollary on p74 of [13], and to part of the corollary to Theorem 9 on p50 of [23].

### 23.11 Reducibility and diagonalizability

Let us return to the same type of situation as at the beginning of the previous section. Thus we let \( k \) be a field, and \((A, [\cdot, \cdot]_A)\) be a finite-dimensional Lie algebra over \( k \). We also let \( B \) be a Lie subalgebra of \( A \) such that every element of \( B \) is ad-diagonalizable as an element of \( A \), which implies that \( B \) is commutative as a Lie algebra, as before. If \( \alpha \) is an element of the dual \( B' \) of \( B \), as a vector space over \( k \), then \( A_\alpha \) is the set of \( x \in A \) such that \([w, x]_A = \alpha(w)x\) for every \( w \in B \). We let \( \Phi_B \) be the set of \( \alpha \in B' \) such that \( \alpha \neq 0 \) and \( A_\alpha \neq \{0\} \), which is a finite set with the property that \( A \) corresponds to the direct sum of \( A_\alpha \), \( \alpha \in \Phi_B \cup \{0\} \), as a vector space over \( k \).

Remember that \([A_\alpha, A_\beta] \subseteq A_{\alpha + \beta}\) for every \( \alpha, \beta \in B' \), as in Section 17.2. Let \( B_0 \) be a subsemigroup of \( B' \), as a commutative group with respect to addition, that contains \( \Phi_B \cup \{0\} \). In particular, we are in the same type of situation as in Section 22.3, with \( \Xi = B_0 \) and \( \Psi = \Phi_B \cup \{0\} \). If \( \Psi_0 \subseteq B_0 \), then let \( A(\Psi_0) \) be the subset of \( A_0 \) consisting of finite sums of elements of \( A_\alpha \), \( \alpha \in \Psi_0 \), as in Section 22.4. If

\[
(23.11.1) \quad (\Psi_0 + \Psi_0) \cap (\Phi_B \cup \{0\}) \subseteq \Psi_0,
\]

then \( A(\Psi_0) \) is a Lie subalgebra of \( A \), as before.

Let \( \hat{B}'_1, \hat{B}'_2 \) be complementary linear subspaces of \( B' \), so that

\[
(23.11.2) \quad \hat{B}'_1 \cap \hat{B}'_2 = \{0\}
\]

and

\[
(23.11.3) \quad \hat{B}'_1 + \hat{B}'_2 = B'.
\]
Suppose that
\[ \Phi_B \subseteq \hat{B}_1' \cup \hat{B}_2', \]
and put
\[ \Phi_{B_j} = \Phi_B \cap \hat{B}_j' \]
for \( j = 1, 2 \). Note that \( \Phi_B = \Phi_{B,1} \cup \Phi_{B,2} \) and
\[ \Phi_{B,1} \cap \Phi_{B,2} = \emptyset, \]
by (23.11.2). If \( \alpha \in \Phi_{B,1} \) and \( \beta \in \Phi_{B,2} \), then
\[ \alpha + \beta \notin \Phi_{B} \cup \{0\}, \]
because \( \alpha + \beta \notin \hat{B}_1' \cup \hat{B}_2' \). It follows that
\[ [A_\alpha, A_\beta] = \{0\} \]
in this case.

We are now in a situation like one considered in Section 22.14, with \( \Phi = \Phi_B \)
and \( \Phi_j = \Phi_{B,j} \) for \( j = 1, 2 \). More precisely, (22.14.7), (22.14.11), (22.14.22),
and (22.14.25) hold under these conditions, as in the preceding paragraph. In
particular, \( A(\Phi_{B,j} \cup \{0\}) \) is a Lie subalgebra of \( A \) for \( j = 1, 2 \), as before. Let \( A^j \)
be the Lie subalgebra of \( A \) generated by \( A(\Phi_{B,j}) \) for \( j = 1, 2 \), as before, which
is contained in \( A(\Phi_{B,j} \cup \{0\}) \). Remember that
\[ [A^1, A^2] = \{0\}, \]
and that \( A^j \) is an ideal in \( A \) for \( j = 1, 2 \), as a Lie algebra over \( k \).

Put
\[ B_1 = \{w \in B : \beta(w) = 0 \text{ for every } \beta \in \hat{B}_2'\}, \]
\[ B_2 = \{w \in B : \beta(w) = 0 \text{ for every } \beta \in \hat{B}_1'\}. \]
These are complementary linear subspaces of \( B \), because \( \hat{B}_1', \hat{B}_2' \) are com-
plementary linear subspaces of \( B' \), by hypothesis. Observe that
\[ [B_1, A_\beta] = \{0\} \]
for every \( \beta \in \Phi_{B,2} \), and similarly that
\[ [B_2, A_\beta] = \{0\} \]
for every \( \beta \in \Phi_{B,1} \), by the definition of \( A_\beta \). Equivalently, this means that
\[ [B_1, A(\Phi_{B,2})] = [B_2, A(\Phi_{B,1})] = \{0\}, \]
which is the same as saying that
\[ A(\Phi_{B,1}) \subseteq C_A(B_2), \quad A(\Phi_{B,2}) \subseteq C_A(B_1). \]
It follows that
\[(23.11.16)\quad A^1 \subseteq C_A(B_2), \quad A^2 \subseteq C_A(B_1),\]
because the centralizer $C_A(E)$ of $E \subseteq A$ is a Lie subalgebra of $A$.

Suppose from now on in this section that $B$ is its own centralizer in $A$, so that $B = A_0$. In particular, this means that $A_0$ is commutative as a Lie algebra over $k$. This implies that $A^1 \cap A^2$ is contained in the center $Z(A)$ of $A$ as a Lie algebra, as in Section 22.14.

If $j \in \{1, 2\}$ and $\alpha \in \Phi_{B,j}$, then
\[(23.11.17)\quad [A_\alpha, A_{-\alpha}] \subseteq A^j \cap A_0 = A^j \cap B,\]
Note that $A_{-\alpha} = \{0\}$ unless $-\alpha \in \Phi_{B,j}$ too. As in Section 22.14, $A^j$ is the same as the set of elements of $A$ that can be expressed as finite sums of elements of $A(\Phi_{B,j})$ and $[A_\alpha, A_{-\alpha}]$ with $\alpha \in \Phi_{B,j}$, because $A_0$ is commutative as a Lie algebra over $k$.

Suppose for the moment that $\alpha \in \Phi_{B,1}$, $\beta \in \Phi_{B,2}$, $x_\alpha \in A_\alpha$, $y_\alpha \in A_{-\alpha}$, and $z_\beta \in A_\beta$. This implies that
\[(23.11.18)\quad [x_\alpha, z_\beta]_A = [y_\alpha, z_\beta]_A = 0,\]
because $\pm \alpha + \beta \notin \hat{B}_1 \cup \hat{B}_2$. It follows that
\[(23.11.19)\quad [[x_\alpha, y_\alpha]_A, z_\beta]_A = 0,\]
by the Jacobi identity. We also have that $[x_\alpha, y_\alpha]_A \in A_0 = B$, so that
\[(23.11.20)\quad [[x_\alpha, y_\alpha]_A, z_\beta]_A = \beta([x_\alpha, y_\alpha]_A) z_\beta,\]
by definition of $A_\beta$. If $z_\beta \neq 0$, then we get that
\[(23.11.21)\quad \beta([x_\alpha, y_\alpha]_A) = 0.\]
Of course, we can always choose $z_\beta$ to be a nonzero element of $A_\beta$, because $\beta \in \Phi_{B,2} \subseteq \Phi_B$. This shows that (23.11.21) holds for every $x_\alpha \in A_\alpha$ and $y_\alpha \in A_{-\alpha}$ when $\alpha \in \Phi_{B,2}$ and $\beta \in \Phi_{B,1}$.

Let us suppose from now on in this section that $Z(A) = \{0\}$. This implies that the linear span of $\Phi_B$ in $\hat{B}'$ is equal to $\hat{B}'$, as in Section 17.5. It follows that the linear span of $\Phi_{B,j}$ in $\hat{B}'_j$ is equal to $\hat{B}'_j$ for $j = 1, 2$ in this situation. This means that (23.11.21) holds for every $x_\alpha \in A_\alpha$ and $y_\alpha \in A_{-\alpha}$ when $\alpha \in \Phi_{B,j}$ and $\beta \in \hat{B}'_j$, and when $\alpha \in \Phi_{B,2}$ and $\beta \in \hat{B}'_1$. This is the same as saying that
\[(23.11.22)\quad [A_\alpha, A_{-\alpha}] \subseteq B_j\]
for $j = 1, 2$ and $\alpha \in \Phi_{B,j}$.

Using (23.11.22), we get that
\[(23.11.23)\quad A^j \subseteq A(\Phi_{B,j}) + B_j\]
for $j = 1, 2$. One can check that $A(\Phi_{B,j}) + B_j$ is a Lie subalgebra of $A$ for $j = 1, 2$, because $B$ is commutative as a Lie subalgebra of $A$, and $[B, A(\Phi_{B,j})] \subseteq A(\Phi_{B,j})$ for $j = 1, 2$. We also have that

$$[A(\Phi_{B,1}) + B_1, A(\Phi_{B,2}) + B_2] = \{0\},$$

because of (23.11.8), (23.11.14), and the fact that $B$ is commutative as a Lie algebra. It is easy to see that

$$A(\Phi_{B,1}) + B_1, \quad A(\Phi_{B,2}) + B_2$$

are complementary linear subspaces of $A$, because $B_1, B_2$ are complementary linear subspaces of $B = A_0$. It follows that these are ideals in $A$, and that $A$ corresponds to their direct sum, as a Lie algebra over $k$.

This is related to the proposition on p73 of [13], and to part of the corollary to Theorem 9 on p50 of [23], as before.
Chapter 24

Strong $\text{ad}$-nilpotence and automorphisms

24.1 Invariant subspaces and nilpotent vectors

Let $k$ be a field, let $V$ be a vector space over $k$, and let $T$ be a linear mapping from $V$ into itself. Remember that

\[(24.1.1) \quad V_0 = \{v \in V : T^l(v) = 0 \text{ for some } l \in \mathbb{Z}_+\}\]

is a linear subspace of $V$, as in Section 10.7. Clearly

\[(24.1.2) \quad T(V_0) \subseteq V_0.\]

In fact,

\[(24.1.3) \quad T^{-1}(V_0) = V_0.\]

More precisely, if $v \in V$ and $T(v) \in V_0$, then it is easy to see that $v \in V_0$.

Let $W$ be a linear subspace of $V$, and suppose that

\[(24.1.4) \quad T(W) \subseteq W.\]

Put

\[(24.1.5) \quad W_0 = W \cap V_0 = \{w \in W : T^l(w) = 0 \text{ for some } l \in \mathbb{Z}_+\},\]

which is a linear subspace of $W$. Of course,

\[(24.1.6) \quad T(W_0) \subseteq W_0,\]

as in (24.1.2). Similarly,

\[(24.1.7) \quad W \cap T^{-1}(W_0) = W_0,\]

as in (24.1.3).
24.2. **FINITE-DIMENSIONAL SPACES**

Let $q$ be the canonical quotient mapping from $V$ onto the quotient vector space $V/W$. As usual, there is a unique linear mapping $T_{V/W}$ from $V/W$ into itself such that

$$T_{V/W} \circ q = q \circ T$$

on $V$. This implies that

$$(T_{V/W})^l \circ q = q \circ T^l$$

on $V$ for every positive integer $l$.

Put

$$(V/W)_0 = \{ z \in V/W : (T_{V/W})^l(z) = 0 \text{ for some } l \in \mathbb{Z}_+ \},$$

which is a linear subspace of $V/W$, as before. If $v \in V$, then

$$(T_{V/W})^l(q(v)) = q(T^l(v))$$

for every positive integer $l$, as in (24.1.9). Thus (24.1.10) consists of the $q(v)$ with $v \in V$ such that (24.1.11) is equal to 0 for some $l \geq 1$. Equivalently, this means that

$$T^l(v) \in W$$

for some $l \geq 1$. In particular,

$$q(V_0) \subseteq (V/W)_0.$$ 

Let $q_0$ be the canonical quotient mapping from $W$ onto the quotient vector space $W/W_0$. As before, there is a unique linear mapping $T_{W/W_0}$ from $W/W_0$ into itself such that

$$T_{W/W_0} \circ q_0 = q_0 \circ T$$

on $W$. Thus

$$(T_{W/W_0})^l \circ q_0 = q_0 \circ T^l$$

on $W$ for every positive integer $l$, as before. Observe that the kernel of $T_{W/W_0}$ is trivial in $W/W_0$, because of (24.1.7).

### 24.2 Finite-dimensional spaces

Let us continue with the same notation and hypotheses as in the previous section. Suppose in addition that the dimension of $W$ is finite, as a vector space over $k$. Of course, this implies that the dimension of $W/W_0$ is finite. Remember that the kernel of $T_{W/W_0}$ is trivial in $W/W_0$. It follows that $T_{W/W_0}$ maps $W/W_0$ onto itself.

If $w \in W$ and $l \in \mathbb{Z}_+$, then we get that there is a $u_l \in W$ such that

$$(T_{W/W_0})^l(q_0(u_l)) = q_0(w).$$

Equivalently, this means that

$$q_0(T^l(u)) = q_0(w).$$
This is the same as saying that

\[(24.2.3) \quad T^i(u_l) - w \in W_0.\]

Let \(v \in V\) be given, and suppose that \(q(v) \in (V/W)_0\). This implies that \(T^i(v) \in W\) for some positive integer \(l\), as in (24.1.12). It follows that there is a \(u_l \in W\) such that

\[(24.2.4) \quad T^i(u_l) - T^i(v) \in W_0,\]

as in the preceding paragraph. Equivalently, this means that

\[(24.2.5) \quad T^i(v - u_l) \in W_0.\]

Hence there is a positive integer \(r\) such that

\[(24.2.6) \quad T^r(T^i(v - u_l)) = 0,\]

by the definition (24.1.5) of \(W_0\).

Thus \(T^{i+r}(v - u_l) = 0\), so that

\[(24.2.7) \quad v - u_l \in V_0.\]

Of course, \(q(v) = q(v - u_l)\), because \(u_l \in W\). It follows that

\[(24.2.8) \quad q(v) = q(v - u_l) \in q(V_0).\]

Combining this with (24.1.13), we get that

\[(24.2.9) \quad q(V_0) = (V/W)_0\]

under these conditions.

Alternatively, suppose that \(V\) has finite dimension, as a vector space over \(k\). It is well known that the characteristic polynomial of \(T\) on \(V\) is equal to the product of the characteristic polynomial of the restriction of \(T\) to \(W\) and the characteristic polynomial of \(T_{V/W}\) on \(V/W\). In particular, the order of vanishing of the characteristic polynomial of \(T\) on \(V\) at 0 is the same as the sum of the orders of vanishing of the characteristic polynomials of the restriction of \(T\) to \(W\) and \(T_{V/W}\) at 0. This means that the dimension of \(V_0\) is equal to the sum of the dimensions of \(W_0\) and \(q(V_0)\), as vector spaces over \(k\). Note that the kernel of the restriction of \(q\) to \(V_0\) is equal to \(W_0\), because \(W_0 = W \cap V_0\), as in (24.1.5). This implies that the dimension of \(V_0\) is equal to the sum of the dimensions of \(W_0\) and \(q(V_0)\) too. Hence the dimension of \(q(V_0)\) is equal to the dimension of \((V/W)_0\). This permits one to obtain (24.2.9) from (24.1.13).

### 24.3 Strong ad-nilpotence

Let \(k\) be a field of characteristic 0, and let \((A, [\cdot, \cdot], \cdot)\) be a finite-dimensional Lie algebra over \(k\). An element \(x\) of \(A\) is said to be **strongly ad-nilpotent** if there are a \(y \in A\), \(\alpha \in k\), and \(l \in \mathbb{Z}_+\) such that \(\alpha \neq 0\) and

\[(24.3.1) \quad (\text{ad}_y - \alpha I)^l(x) = 0,\]
where $I$ is the identity mapping on $A$. This definition is given on p82 of [13], where $k$ is also taken to be algebraically closed. In that case, if $x$ is strongly ad-nilpotent, then $x$ is ad-nilpotent in $A$, by the remarks in Section 18.1, as in [13].

Otherwise, let $k_1$ be an algebraically closed field that contains $k$. We may as well identify $A$ with $k^n$ for some positive integer $n$ as a vector space over $k_1$, using a basis for $A$. Let $A_1$ be $k_1^n$, as a vector space over $k_1$. We can extend the Lie bracket on $A$ to $A_1$, so that $A_1$ becomes a Lie algebra over $k_1$. If $x \in A$ is strongly ad-nilpotent as an element of $A$, then $x$ is strongly ad-nilpotent as an element of $A_1$, as before. This implies that $x$ is ad-nilpotent as an element of $A$ as well. Another approach to this will be given in Section 24.5.

If $A$ is semisimple, and $x \in A$ is ad-nilpotent, then $x$ is strongly ad-nilpotent, as in Sections 14.1 and 14.2.

Let $N(A)$ be the set of $x \in A$ such that $x$ is strongly ad-nilpotent in $A$. As a subgroup of $\text{Int}(A)$, $N(A)$ is invariant under Lie algebra automorphisms of $A$.

If $x \in A$ is ad-nilpotent in $A$, then the exponential $\exp \text{ad}_x$ of $\text{ad}_x$ on $A$ defines a Lie algebra automorphism of $A$, as in Section 14.11. Remember that $\text{Int}(A)$ denotes the subgroup of the automorphism group of $A$ generated by these automorphisms. Let $\mathcal{E}(A)$ be the subgroup of $\text{Int}(A)$ generated by exponentials of strongly ad-nilpotent elements of $A$. One can check that $\mathcal{E}(A)$ is a normal subgroup of the group of all Lie algebra automorphisms of $A$, because $N(A)$ is invariant under Lie algebra automorphisms of $A$, as on p82 of [13]. If $A$ is semisimple, then $\mathcal{E}(A) = \text{Int}(A)$, because $N(A)$ contains all ad-nilpotent elements of $A$, as before.

Let $B$ be a Lie subalgebra of $A$, and let $x$ be an element of $B$. Thus $\text{ad}_{B,x}$ is the same as the restriction of $\text{ad}_{A,x}$ to $B$. Suppose that $x \in N(B)$, so that $x \in N(A)$, and $\text{ad}_{A,x}$ is nilpotent on $A$. Note that $\exp \text{ad}_{B,x}$ is the same as the restriction of $\exp \text{ad}_{A,x}$ to $B$.

Let $\mathcal{E}(A,B)$ be the subgroup of $\mathcal{E}(A)$ generated by elements of the form $\exp \text{ad}_{A,x}$ with $x \in N(B)$. The restriction of every element of $\mathcal{E}(A,B)$ to $B$ is an element of $\mathcal{E}(B)$. It is easy to see that every element of $\mathcal{E}(B)$ can be obtained from an element of $\mathcal{E}(A,B)$ in this way. This corresponds to some remarks on p82 of [13].

### 24.4 Strong ad-nilpotence and homomorphisms

Let $k$ be a field of characteristic 0, and let $(A, [\cdot ,\cdot ]_A)$, $(C, [\cdot ,\cdot ]_C)$ be finite-dimensional Lie algebras over $k$. Also let $\psi$ be a Lie algebra homomorphism from $A$ into $C$. If $y \in A$, then

\begin{equation}
\psi \circ \text{ad}_{A,y} = \text{ad}_{C,\psi(y)} \circ \psi,
\end{equation}
as in Section 14.8. This implies that
\[ \psi \circ (\text{ad}_{A,\alpha I_A} - \alpha I_A) = (\text{ad}_{C,\psi(y)} - \alpha I_C) \circ \psi \]
for every \( \alpha \in k \), where \( I_A, I_C \) are the identity mappings on \( A, C \), respectively.
It follows that
\[ \psi \circ (\text{ad}_{A,\alpha I_A} - \alpha I_A)^l = (\text{ad}_{C,\psi(y)} - \alpha I_C)^l \circ \psi \]
for every \( l \in \mathbb{Z}_+ \).

Suppose that \( x \in A \) is strongly ad-nilpotent in \( A \), so that there are \( y \in A \), \( \alpha \in k \), and \( l \in \mathbb{Z}_+ \) such that \( \alpha \neq 0 \) and
\[ (\text{ad}_{A,\alpha I_A} - \alpha I_A)^l(x) = 0; \]
as in (24.4.1). This implies that
\[ (\text{ad}_{C,\psi(y)} - \alpha I_C)^l(\psi(x)) = 0, \]
by (24.4.3). This means that \( \psi(x) \) is strongly ad-nilpotent in \( C \), so that
\[ \psi(N(A)) \subseteq N(C). \]
In particular, \( x \) and \( \psi(x) \) are ad-nilpotent in \( A \) and \( C \), respectively, as before.

Let \( y \in A \) and \( \alpha \in k \) be given again, and observe that
\[ \{ x \in A : (\text{ad}_{A,\alpha I_A} - \alpha I_A)^l(x) = 0 \text{ for some } l \in \mathbb{Z}_+ \} \]
is a linear subspace of \( A \). This corresponds to (24.1.1), with \( V = A \) and \( T = \text{ad}_{A,\alpha I_A} - \alpha I_A \). Similarly,
\[ \{ z \in C : (\text{ad}_{C,\psi(y)} - \alpha I_C)^l(z) = 0 \text{ for some } l \in \mathbb{Z}_+ \} \]
is a linear subspace of \( C \). Clearly \( \psi \) maps (24.4.8) into (24.4.9), because of (24.4.3). If, in the notation of Section 24.1, we take \( W \) to be the kernel of \( \psi \), then (24.1.4) holds, because the kernel of \( \psi \) is an ideal in \( A \).

Suppose now that \( \psi \) maps \( A \) onto \( C \), so that \( V/W \) corresponds to \( C \), and the quotient mapping \( q \) from \( V \) onto \( V/W \) corresponds to \( \psi \). Note that the linear mapping \( T_{V/W} \) induced on \( V/W \) by \( T \) as in Section 24.1 corresponds to \( \text{ad}_{C,\psi(y)} - \alpha I_C \), and that (24.1.10) corresponds to (24.4.9). In this case, \( \psi \) maps (24.4.8) onto (24.4.9), as in (24.2.9). It follows that
\[ (\text{ad}_{A,\alpha I_A} - \alpha I_A)^l(x) = 0, \]
under these conditions. This corresponds to some remarks on p82 of [13].

If \( \sigma_C \in \mathcal{E}(C) \), then there is a \( \sigma_A \in \mathcal{E}(A) \) such that
\[ \psi \circ \sigma_A = \sigma_C \circ \psi. \]
This can be obtained from (24.4.7) and (24.4.10). This corresponds to the lemma on p82 of [13].
24.5 Another criterion for nilpotence

Let \( k \) be a field of characteristic 0, let \( \mathcal{A} \) be an associative algebra over \( k \), and let \( \delta \) be a derivation on \( \mathcal{A} \). If \( \lambda \in k \), then put

\[
E_\lambda(\delta) = \{ a \in \mathcal{A} : (\delta - \lambda I)^l(a) = 0 \text{ for some } l \in \mathbb{Z}_+ \},
\]

as in Sections 10.7 and 10.9, where \( I \) is the identity mapping on \( \mathcal{A} \). If \( b \in E_\lambda(\delta) \) and \( c \in E_\mu(\delta) \) for some \( \mu \in k \), then \( bc \in E_{\lambda+\mu}(\delta) \), as in Section 10.9. It follows that

\[
b^j \in E_{j\lambda}(\delta)
\]

for every positive integer \( j \).

Suppose that \( \lambda \neq 0 \), so that the elements of \( k \) of the form \( j \cdot \lambda \), \( j \in \mathbb{Z}_+ \), are distinct, because \( k \) has characteristic 0. If, for each \( j \in \mathbb{Z}_+ \), we have that \( b^j \neq 0 \), then the \( b^j \)'s are linearly independent in \( \mathcal{A} \), as in Section 10.7. If \( \mathcal{A} \) has finite dimension as a vector space over \( k \), then it follows that \( b^j = 0 \) for some \( j \geq 1 \). This extends the remarks at the beginning of Section 14.2.

If \( a \in \mathcal{A} \), then \( \delta_\lambda(x) = [a, x] = ax - xa \) defines a derivation on \( \mathcal{A} \). Thus if \( b \in E_\lambda(\delta) \) for some nonzero \( \lambda \in k \) and \( \mathcal{A} \) has finite dimension as a vector space over \( k \), then \( b \) is nilpotent in \( \mathcal{A} \).

Remember that \( \mathcal{A} \) may be considered as a Lie algebra over \( k \), with respect to the commutator bracket corresponding to multiplication in \( \mathcal{A} \). If \( b \in \mathcal{A} \) is strongly ad-nilpotent in \( \mathcal{A} \) as a Lie algebra, then \( b \) is nilpotent in \( \mathcal{A} \), as in the preceding paragraph.

Let \( (\mathcal{A}, [\cdot, \cdot]_A) \) be a finite-dimensional Lie algebra over \( k \). If \( x \in \mathcal{A} \) is strongly ad-nilpotent, then \( \text{ad}_x \) is strongly ad-nilpotent as an element of the Lie algebra of linear mappings from \( \mathcal{A} \) into itself. This follows from (24.4.6), using the adjoint representation of \( \mathcal{A} \). This implies that \( \text{ad}_x \) is nilpotent as a linear mapping on \( \mathcal{A} \), as in the previous paragraph.

24.6 Nilpotence and automorphisms

Let \( k \) be a field, and let \( f(t) = \sum_{j=0}^n f_j t^j \) be a polynomial function of \( t \in k \) with coefficients in a vector space over \( k \). If \( f_j \neq 0 \) for some \( j \), then \( f(t) = 0 \) for at most \( n \) \( t \in k \), by standard arguments.

Suppose from now on in this section that \( k \) has characteristic 0. Let \( \mathcal{A} \) be an associative algebra over \( k \) with a multiplicative identity element \( e \), and let \( a \) be a nilpotent element of \( \mathcal{A} \). Thus \( \exp(ta) \) defines a polynomial function of \( t \in k \) with values in \( \mathcal{A} \). If \( a \neq 0 \), then \( \exp(ta) = e \) for only finitely many \( t \in k \), as in the previous paragraph.

Let \( (\mathcal{A}, [\cdot, \cdot]_A) \) be a Lie algebra over \( k \). If \( x \in \mathcal{A} \) is ad-nilpotent, then \( \exp(t \, \text{ad}_x) \) defines a polynomial function of \( t \in k \) with values in the algebra of linear mappings from \( \mathcal{A} \) into itself. If \( \text{ad}_x \neq 0 \) as a linear mapping from \( \mathcal{A} \) into itself, then \( \exp(t \, \text{ad}_x) \) is equal to the identity mapping on \( \mathcal{A} \) for only finitely many \( t \in k \), as before.
Let $y \in A$ be given, and for each $\alpha \in k$, let $A_{y,\alpha}$ be the set of $x \in A$ such that $(\text{ad}_y - \alpha I)^l(x) = 0$ for some positive integer $l$, as in Section 18.1. Remember that $A_{y,\alpha} \neq \{0\}$ exactly when $\alpha$ is an eigenvalue of $\text{ad}_y$ on $A$. In particular, if $y$ is ad-nilpotent as an element of $A$, then $A_{y,\alpha} = \{0\}$ when $\alpha \neq 0$.

If $A$ is nilpotent as a Lie algebra over $k$, then every $y \in A$ is ad-nilpotent, and it follows that $0$ is the only element of $A$ that is strongly ad-nilpotent. This implies that $E(A)$ contains only the identity mapping on $A$, as in Exercise 1 on p87 of [13].

Suppose now that $k$ is algebraically closed, and that $A$ has finite dimension, as a vector space over $k$. If $y \in A$, then $A$ corresponds to the direct sum of the linear subspaces $A_{y,\alpha}$, where $\alpha \in k$ is an eigenvalue of $\text{ad}_y$ on $A$, as a vector space over $k$, as in Section 18.1.

If $\text{ad}_y$ is not nilpotent on $A$, then it follows that $\text{ad}_y$ has a nonzero eigenvalue $\alpha \in k$. This means that there is an $x \in A$ such that $x \neq 0$ and $\text{ad}_y(x) = \alpha x$, so that $x$ is strongly ad-nilpotent in particular. Note that $\text{ad}_x(y) = -\text{ad}_y(x) = -\alpha x \neq 0$, so that $\text{ad}_x \neq 0$. Of course, $t \text{ad}_x$ is strongly ad-nilpotent for every $t \in k$ too, so that

\begin{equation}
(24.6.1) \quad \exp(t \text{ad}_x) \in E(A)
\end{equation}

for every $t \in k$.

If $A$ is not nilpotent as a Lie algebra, then there is a $y \in A$ that is not ad-nilpotent, by Engel’s theorem. Under these conditions, we get that $E(A)$ is nontrivial, as in Exercise 1 on p87 of [13].

### 24.7 Conjugacy and solvability

Let $k$ be an algebraically closed field of characteristic 0, and let $(A, [\cdot, \cdot]_A)$ be a finite-dimensional solvable Lie algebra over $k$. If $C_1$, $C_2$ are Cartan subalgebras of $A$, then there is an element of the group $E(A)$ defined in Section 24.3 that maps $C_1$ onto $C_2$. This is the theorem on p82 of [13], which is related to Theorem 2 on p12 of [23].

Of course, if $A$ is nilpotent as a Lie algebra, then $A$ may be considered as a Cartan subalgebra of itself. More precisely, $A$ is the only Cartan subalgebra of itself in this case, because Cartan subalgebras are maximal among nilpotent Lie subalgebras, as mentioned at the beginning of Section 18.11.

To prove the theorem, one can use induction on the dimension of $A$. We may also suppose that $A$ is not nilpotent as a Lie algebra over $k$, as in the preceding paragraph. In particular, this means that the dimension of $A$ is at least two.

Because $A \neq \{0\}$ is solvable as a Lie algebra, $A$ has nontrivial ideals that are commutative as Lie algebras. Let $A_0$ be a nontrivial ideal in $A$ that is commutative as a Lie algebra, and with minimal dimension among such ideals.

Let $q_0$ be the natural quotient mapping from $A$ onto $A/A_0$, and note that $A/A_0$ is a solvable Lie algebra over $k$, with dimension strictly less than the dimension of $A$. Under these conditions, $q_0(C_1)$ and $q_0(C_2)$ are Cartan subalgebras of $A/A_0$, as in Section 18.11.
24.8. **THE CASE WHERE** \( A = B_2 \)

Our induction hypothesis implies that there is an element \( \sigma_0 \) of \( \mathcal{E}(A/A_0) \) such that

\[
\sigma_0(q_0(C_1)) = q_0(C_2).
\]

This leads to an element \( \sigma \) of \( \mathcal{E}(A) \) such that

\[
q_0 \circ \sigma = \sigma_0 \circ q_0,
\]

as in Section 24.4.

Put

\[
B_1 = q_0^{-1}(q_0(C_1)), \quad B_2 = q_0^{-1}(q_0(C_2)).
\]

These are Lie subalgebras of \( A \), with

\[
C_1 \subseteq B_1, \quad C_2 \subseteq B_2.
\]

It is easy to see that

\[
\sigma(B_1) = B_2,
\]

using (24.7.1) and (24.7.2).

Note that \( C_1, C_2 \) may be considered as Cartan subalgebras of \( B_1, B_2 \), respectively, as in Section 18.11. It follows that \( \sigma(C_1) \) may be considered as a Cartan subalgebra of \( B_2 \), by (24.7.5).

Suppose for the moment that

\[
B_2 \neq A.
\]

In this case, our induction hypothesis implies that there is an element \( \tau_2 \) of \( \mathcal{E}(B_2) \) such that

\[
\tau_2(\sigma(C_1)) = C_2.
\]

As in Section 24.3, there is an element \( \tau \) of \( \mathcal{E}(A, B_2) \subseteq \mathcal{E}(A) \) whose restriction to \( B_2 \) is equal to \( \tau_2 \). In particular, this implies that

\[
\tau(\sigma(C_1)) = C_2.
\]

This completes the proof of the theorem in this case, because \( \tau \circ \sigma \in \mathcal{E}(A) \).

24.8 **The case where** \( A = B_2 \)

Let us continue with the same notation and hypotheses as in the previous section, except that we suppose now that \( A = B_2 \). In this case, \( \sigma(B_1) = B_2 = A \), so that \( B_1 = A \) as well. Equivalently, this means that

\[
C_1 + A_0 = C_2 + A_0 = A.
\]

If \( x \in A \) and \( \alpha \in k \), then let \( A_{x,\alpha} \) be the set of \( y \in A \) such that

\[
(\text{ad}_x - \alpha I)^l(y) = 0
\]
for some positive integer $l$, as in Section 18.1, and where $I$ is the identity mapping on $A$. This is a linear subspace of $A$, and $A$ corresponds to the direct sum of $A_{x,\alpha}$, as a vector space over $k$, where $\alpha \in k$ is an eigenvalue of $\text{ad}_x$ on $A$, as before, because $k$ is algebraically closed, by hypothesis.

Remember that $A_0$ is an ideal in $A$, so that $\text{ad}_x(A_0) \subseteq A_0$. If $\alpha \in k$, then $A_0 \cap A_{x,\alpha}$ consists of $y \in A_0$ such that (24.8.2) holds for some $l \in \mathbb{Z}_+$, which is the analogue of $A_{x,\alpha}$ for the restriction of $\text{ad}_x$ to $A_0$. As before, $A_0$ corresponds to the direct sum of $A_0 \cap A_{x,\alpha}$, where $\alpha \in k$ is an eigenvalue of $\text{ad}_x$ on $A_0$, as a vector space over $k$.

Remember that $[A_{x,\alpha}, A_{x,\beta}] \subseteq A_{x,\alpha + \beta}$ for every $\alpha, \beta \in k$, as in Section 18.1. This implies that

\[(24.8.3) \quad [A_{x,0}, A_0 \cap A_{x,\alpha}] \subseteq A_0 \cap A_{x,\alpha} \]

for every $\alpha \in k$, because $A_0$ is an ideal in $A$.

Because $C_2$ is a Cartan subalgebra of $A$, there is an $x \in A$ such that

\[(24.8.4) \quad C_2 = A_{x,0}, \]

as in Section 18.8. In this case, one can check that

\[(24.8.5) \quad A_0 \cap A_{x,\alpha} \text{ is an ideal in } A \]

for every $\alpha \in k$. More precisely, this uses (24.8.1), (24.8.3), and the fact that $A_0$ is commutative as a Lie subalgebra of $A$.

If

\[(24.8.6) \quad A_0 \cap A_{x,\alpha} \neq \{0\} \]

for some $\alpha \in k$, then

\[(24.8.7) \quad A_0 \cap A_{x,\alpha} = A_0, \]

because $A_0$ is supposed to be minimal among nontrivial ideals that are commutative as Lie subalgebras of $A$. Note that (24.8.6) holds for some $\alpha \in k$, because $A_0 \neq \{0\}$, and $A_0$ corresponds to the direct sum of $A_0 \cap A_{x,\alpha}$ over some $\alpha \in k$, as a vector space over $k$.

If (24.8.7) holds with $\alpha = 0$, then $A_0 \subseteq A_{x,0} = C_2$, which means that $A = C_2$, by (24.8.1). This would imply that $A$ is nilpotent as a Lie algebra, by definition of a Cartan subalgebra. However, $A$ is not supposed to be nilpotent, as in the previous section, and so we get that $A_0 \cap A_{x,0} = \{0\}$.

It follows that (24.8.7) holds for some nonzero $\alpha \in k$, so that $A_0 \subseteq A_{x,\alpha}$. This implies that

\[(24.8.8) \quad A_0 = A_{x,\alpha}, \]

because of (24.8.1), (24.8.4), and the fact that $A$ corresponds to the direct sum of $A_{x,\beta}$ over some $\beta \in k$, as a vector space over $k$.

Using (24.8.1), we get that $x$ can be expressed as

\[(24.8.9) \quad x = y + z, \]

with $y \in C_1$ and $z \in A_0$. More precisely, $z \in A_{x,\alpha}$, by (24.8.8).
24.8. \textit{THE CASE WHERE }$A = B_2$

It is easy to see that $\text{ad}_x(A_{x,\alpha}) \subseteq A_{x,\alpha}$, by the definition of $A_{x,\alpha}$. Because $z \in A_{x,\alpha}$ and $\alpha \neq 0$, one can express $z$ as

\begin{equation}
(24.8.10)
z = \text{ad}_x(w),
\end{equation}

where $w$ is a linear combination of $(\text{ad}_x)^j(z)$, for finitely many positive integers $j$. In particular, this implies that $w \in A_{x,\alpha}$.

Equivalently, $w \in A_0$, so that $\text{ad}_w(A) \subseteq A_0$, because $A_0$ is an ideal in $A$. This implies that

\begin{equation}
(24.8.11)
\text{ad}_w \circ \text{ad}_w = 0
\end{equation}
on $A$, because $A_0$ is commutative as a Lie subalgebra of $A$.

Thus

\begin{equation}
(24.8.12)
\exp(\text{ad}_w) = I + \text{ad}_w,
\end{equation}

so that

\begin{equation}
(24.8.13)
\exp(\text{ad}_w)(x) = x + \text{ad}_w(x) = x - \text{ad}_x(w) = x - z = y.
\end{equation}

Put

\begin{equation}
(24.8.14)
C = (\exp(\text{ad}_w))(C_2),
\end{equation}

which is a Cartan subalgebra of $A$, because $C_2$ is a Cartan subalgebra of $A$, and $\exp(\text{ad}_w)$ is an automorphism of $A$. One can check that

\begin{equation}
(24.8.15)
C = A_{y,0},
\end{equation}

using (24.8.4) and (24.8.13).

Remember that $y \in C_1$, by construction, so that $\text{ad}_y$ maps $C_1$ into itself. More precisely, $\text{ad}_y$ is nilpotent on $C_1$, because $C_1$ is a Cartan subalgebra of $A$, and thus nilpotent as a Lie algebra. It follows that

\begin{equation}
(24.8.16)
C_1 \subseteq C,
\end{equation}

by (24.8.15).

This implies that

\begin{equation}
(24.8.17)
C_1 = C,
\end{equation}

by the characterization of Cartan subalgebras as minimal Engel subalgebras, as in Section 18.8. Alternatively, one can use the fact that Cartan subalgebras are maximal among nilpotent Lie subalgebras of $A$, as in Section 18.11.

Remember that $w \in A_{x,\alpha}$, so that $w$ is strongly ad-nilpotent in $A$, and $\exp(\text{ad}_w) \in \mathcal{E}(A)$. It follows that there is an element of $\mathcal{E}(A)$ that maps $C_1$ onto $C_2$, by (24.8.14) and (24.8.17). This completes the proof of the theorem stated at the beginning of the previous section.
24.9 Conjugacy in arbitrary Lie algebras

Let \( k \) be an algebraically closed field of characteristic 0 again, and let \( (A, [\cdot, \cdot]_A) \) be a finite-dimensional Lie algebra over \( k \). If \( B_1, B_2 \) are Borel subalgebras of \( A \), then the theorem on p84 of [13] states that

\[(24.9.1) \text{ there is an element of } E(A) \text{ that maps } B_1 \text{ onto } B_2.\]

This is closely related to Theorem 5 on p48 of [23].

If \( C_1, C_2 \) are Cartan subalgebras of \( A \), then the corollary on p84 states that

\[(24.9.2) \text{ there is an element of } E(A) \text{ that maps } C_1 \text{ onto } C_2.\]

This is closely related to Theorem 2 on p12 of [23].

To obtain the corollary from the theorem, remember that Cartan subalgebras are nilpotent, and solvable in particular. This implies that there are Borel subalgebras \( B_1, B_2 \) of \( A \) that contain \( C_1, C_2 \), respectively. Using the theorem, we get \( \sigma \in E(A) \) such that \( \sigma(B_1) = B_2 \). It follows that

\[(24.9.3) \sigma(C_1) \subseteq B_2.\]

Of course, \( \sigma(C_1) \) is a Cartan subalgebra of \( A \) as well. As in Section 18.11, \( \sigma(C_1) \) and \( C_2 \) may also be considered as Cartan subalgebras of \( B_2 \). Because \( B_2 \) is solvable as a Lie algebra over \( k \), there is an element of \( E(B_2) \) that maps \( \sigma(C_1) \) onto \( C_2 \), as in Section 24.7.

Every element of \( E(B_2) \) can be obtained as the restriction of an element of \( E(A, B_2) \subseteq E(A) \) to \( B_2 \), as in Section 24.3. This implies that there is an element of \( E(A) \) that maps \( \sigma(C_1) \) onto \( C_2 \). It follows that there is an element of \( E(A) \) that maps \( C_1 \) onto \( C_2 \), as desired.

24.9.1 The beginning of the proof of the theorem

Let us now begin the proof of the theorem. Note that the theorem holds trivially when \( A \) is solvable as a Lie algebra, in which case \( A \) is the only Borel subalgebra of itself.

To prove the theorem, one uses induction on the dimension of \( A \) as a vector space over \( k \). If the dimension of \( A \) is one, then the theorem is trivial, as in the preceding paragraph.

Remember that \( \text{Rad } A \) denotes the solvable radical of \( A \), as in Section 9.4. If \( B_1, B_2 \) are Borel subalgebras of \( A \), then

\[(24.9.4) \text{Rad } A \subseteq B_1, B_2,\]

as in Section 22.7.

Remember that \( \text{Rad } A \) is an ideal in \( A \), and let \( q \) be the natural quotient mapping from \( A \) onto \( A/\text{Rad}(A) \). Note that \( q(B_1), q(B_2) \) are Borel subalgebras of \( A/\text{Rad } A \), as in Section 22.7. We also have that

\[(24.9.5) B_j = q^{-1}(q(B_j)) \quad \text{for } j = 1, 2,\]
24.9. CONJUGACY IN ARBITRARY LIE ALGEBRAS

If $\text{Rad} A \neq \{0\}$, then the dimension of $A/\text{Rad} A$ is less than the dimension of $A$, and our induction hypothesis implies that there is an element $\sigma$ of $\mathcal{E}(A/\text{Rad} A)$ such that

$\sigma(q(B_1)) = q(B_2)$.  

This leads to an element $\sigma_A$ of $\mathcal{E}(A)$ such that

$q \circ \sigma_A = \sigma \circ q$,  

as in Section 24.4.

It is easy to see that $A\Phi(B_1) = B_2$; using (24.9.5), (24.9.6), and (24.9.7). This completes the proof when $\text{Rad} A \neq \{0\}$.

Thus we may suppose from now on that $\text{Rad} A = \{0\}$, so that $A$ is semisimple as a Lie algebra over $k$. Of course, we may suppose that $A \neq \{0\}$ too.

24.9.2 Standard Borel subalgebras of $A$

Let $A_0$ be a toral subalgebra of $A$, as in Section 17.1, so that every element of $A_0$ is ad-diagonalizable as an element of $A$. Remember that $A_0$ is commutative as a Lie algebra, as before. More precisely, let $A_0$ be a toral subalgebra of $A$ that is maximal with respect to inclusion. Note that $A_0 \neq \{0\}$, as in Section 17.1, because $A \neq \{0\}$, and using our hypotheses on $k$. We also have that the centralizer of $A_0$ in $A$ is equal to $A_0$, as in Section 17.4.

Let $A'_0$ be the dual of $A_0$, as a vector space over $k$, as usual. If $\alpha \in A'_0$, then let $A_\alpha$ be the set of $x \in A$ such that $\text{ad}_w(x) = \alpha(w)x$ for every $w \in A_0$. This is the same as $A_0$ when $\alpha = 0$, because $A_0$ is its own centralizer in $A$, as in the preceding paragraph.

Let $\Phi = \Phi_{A_0}$ be the set of $\alpha \in A'_0$ such that $\alpha \neq 0$ and $A_\alpha \neq \{0\}$. Remember that $A$ corresponds to the direct sum of the subspaces $A_\alpha$, $\alpha \in \Phi \cup \{0\}$, as a vector space over $k$, as in Section 17.2. If $\alpha, \beta \in A'_0$, then $[A_\alpha, A_\beta] \subseteq A_{\alpha + \beta}$, as before.

If $\alpha \in \Phi$, then $A_\alpha$ has dimension one as a vector space over $k$, as in Section 17.7. We may choose $x_\alpha \in A_\alpha$ and $y_\alpha \in A_{-\alpha}$ so that $h_\alpha = [x_\alpha, y_\alpha]$ satisfies $\alpha(h_\alpha) = 2$, as in Section 17.6.

Let $E_Q$ be the linear subspace of $A'_0$, as a vector space over $Q$, spanned by $\Phi$, as in Section 17.12. Let $E_R$ be the corresponding vector space over $R$, as in Section 17.13. Remember that $\Phi$ is a reduced root system in $E_R$, as in Section 22.1.

If $\Psi_0 \subseteq \Phi \cup \{0\}$, then let $A(\Psi_0)$ be the linear subspace of $A$ spanned by $A_\alpha$, $\alpha \in \Psi_0$, as in Section 22.4. Let $\Delta$ be a base for $\Phi$, as a root system in $E_R$, and let $\Phi^+ = \Phi^{\Delta^+}$ be the corresponding set of positive roots, which are the elements of $\Phi$ that can be expressed as linear combinations of elements of $\Delta$ with nonnegative coefficients. Remember that

$B_\Delta = A(\Phi^+ \cup \{0\})$. 

(24.9.9)
is a Borel subalgebra of $A$, as in Section 22.12, which is the standard Borel subalgebra associated to $A_0$ and $\Delta$.

Let $B$ be any other Borel subalgebra of $A$. It suffices to show that there is an element of $\mathcal{E}(A)$ that maps $B_\Delta$ onto $B$.

If $B_\Delta \subseteq B$, then $B_\Delta = B$, because $B_\Delta$ is maximal among solvable Lie subalgebras of $A$. Of course, the problem is trivial in this case.

24.9.3 The second induction hypothesis

To continue the proof, we also use induction on the dimension of $B_\Delta \cap B$, as a vector space over $k$, as on p84 of [13]. If the dimension is equal to the dimension of $B_\Delta$, then $B_\Delta \cap B = B_\Delta$, so that $B_\Delta \subseteq B$, and thus $B_\Delta = B$, as before. If $\tilde{B}$ is a Borel subalgebra of $A$ such that

\[(24.9.10) \quad \dim(B_\Delta \cap \tilde{B}) > \dim B_\Delta \cap B,\]

then we suppose from now on that there is an element of $\mathcal{E}(A)$ that maps $B_\Delta$ onto $\tilde{B}$.

Of course, $B_\Delta$ depends on the choice of the maximal toral subalgebra $A_0$ of $A$, as well as the base $\Delta$ for the corresponding root system $\Phi$, as in the previous subsection. In Section 24.12, we shall consider another maximal toral subalgebra of $A$. Thus $A_0$ should be considered as an arbitrary maximal toral subalgebra of $A$ here, and $\Delta$ as an arbitrary base for the corresponding root system $\Phi$.

As in [13], we consider separately the cases where $B_\Delta \cap B$ is nontrivial or equal to $\{0\}$. In the first case, we consider the subcases where $B_\Delta \cap B$ does or does not have nonzero elements that are ad-nilpotent in $A$. These two subcases are discussed in the next two sections.

24.10 The first subcase

Let us continue with the same notation and hypotheses as in the previous section. In this section, we suppose that

\[(24.10.1) \quad B_\Delta \cap B \neq \{0\},\]

and that

\[(24.10.2) \quad B_\Delta \cap B \text{ has a nonzero element that is ad-nilpotent in } A.\]

This corresponds to Case (i) of (1) on p85 of [13].

We may as well suppose too that

\[(24.10.3) \quad B_\Delta \cap B \neq B_\Delta, B.\]

Otherwise, we would have $B_\Delta \subseteq B$ or $B \subseteq B_\Delta$, which would imply that $B_\Delta = B$. 

Put
\[(24.10.4) \quad N_\Delta = A(\Phi^+), \]
so that \(N_\Delta \subseteq B_\Delta\). Remember that \(N_\Delta\) is a Lie subalgebra of \(A\) that is nilpotent as a Lie algebra over \(k\), and that the elements of \(N_\Delta\) are ad-nilpotent in \(A\), as in Section 22.6.

If \(\alpha \in A_0'\) and \(\alpha \neq 0\), then it is easy to see that \([A_0, A_\alpha] = A_\alpha\). This implies that
\[(24.10.5) \quad [B_\Delta, B_\Delta] = N_\Delta, \]
as in Section 22.6.

We would like to check that
\[(24.10.6) \quad N_\Delta = \{ v \in B_\Delta : \text{ad}_v \text{ is nilpotent on } A \}. \]
We have seen that \(N_\Delta\) is contained in the right side, and so we only have to verify that the right side is contained in \(N_\Delta\). Note that \(B_\Delta\) corresponds to the direct sum of \(A_0\) and \(N_\Delta\), as a vector space over \(k\).

Let \(v \in B_\Delta\) be given, so that \(v\) can be expressed as
\[(24.10.7) \quad v = w + u, \]
with \(w \in A_0\) and \(u \in N_\Delta\). Suppose that \(v \notin N_\Delta\), so that \(w \neq 0\). We would like to check that \(\text{ad}_v\) is not nilpotent on \(A\).

Remember that the linear span of \(\Phi\) in \(A_0'\), as a vector space over \(k\), is equal to \(A_0'\), as in Section 17.5. This implies that the linear span of \(\Delta\) in \(A_0'\) is equal to \(A_0'\), because every element of \(\Phi\) can be expressed as a linear combination of elements of \(\Delta\) with coefficients in \(\mathbb{Z}\). It follows that \(\alpha(w) \neq 0\) for some \(\alpha \in \Delta\), because \(w\) is a nonzero element of \(A_0\).

Observe that \(\text{ad}_v\) maps \(N_\Delta\) into itself, and that \(N_\Delta\) corresponds to the direct sum of \(A_\beta, \beta \in \Phi^+\), as a vector space over \(k\). If \(z \in N_\Delta\), then the component of \(\text{ad}_v(z)\) in \(A_\alpha\) is equal to 0. This implies that the component of \(\text{ad}_v(z)\) in \(A_\alpha\) is the same as the component of \(\text{ad}_w(z)\) in \(A_\alpha\), which is \(\alpha(w)\) times the component of \(z\) in \(A_\alpha\). One can use this to get that \(\text{ad}_v\) is not nilpotent on \(N_\Delta\), because \(A_\alpha \neq \{0\}\).

### 24.10.1 Nilpotent elements of \(B_\Delta \cap B\)

Let \(N\) be the set of elements of \(B_\Delta \cap B\) that are ad-nilpotent in \(A\), so that \(N \neq \{0\}\), by hypothesis. Observe that
\[(24.10.8) \quad N = N_\Delta \cap B, \]
by (24.10.6).

Thus \(N\) is a linear subspace of \(A\), and in fact a Lie subalgebra of \(A\). We also have that
\[(24.10.9) \quad [B_\Delta \cap B, B_\Delta \cap B] \subseteq [B_\Delta, B_\Delta] \cap B = N_\Delta \cap B = N. \]
In particular, \(N\) is an ideal in \(B_\Delta \cap B\), as a Lie algebra over \(k\).
CHAPTER 24. STRONG AD-NILPOTENCE AND AUTOMORPHISMS

Of course, \( x \mapsto \text{ad}_x = \text{ad}_{A_x} \) defines a representation of \( B \) on \( A \), as a vector space over \( k \). If \( x \in [B, B] \), then \( \text{ad}_x \) is nilpotent on \( A \), because \( B \) is solvable, as in Section 14.14. This implies that

\[
(24.10.10) \quad B_\Delta \cap [B, B] \subseteq N.
\]

Note that \( N \) is nilpotent as a Lie algebra over \( k \), because \( N_\Delta \) is nilpotent. This means that \( N \) is solvable as a Lie algebra over \( k \) in particular. It follows that \( N \) is not an ideal in \( A \), because \( A \) is semisimple, and \( N \neq \{0\} \).

Let \( M \) be the normalizer of \( N \) in \( A \), so that \( M \) is a proper Lie subalgebra of \( A \). Clearly

\[
(24.10.11) \quad B_\Delta \cap B \subseteq M,
\]

because \( N \) is an ideal in \( B_\Delta \cap B \).

Let us check that

\[
(24.10.12) \quad B_\Delta \cap B \neq B_\Delta \cap M.
\]

Of course, \( B_\Delta \cap B \) is a linear subspace of \( B_\Delta \), so that the quotient

\[
(24.10.13) \quad B_\Delta / (B_\Delta \cap B)
\]

may be considered as a vector space over \( k \). If \( x \in N \), then \( \text{ad}_x \) maps \( B_\Delta \) and \( B_\Delta \cap B \) into themselves, because \( x \in B_\Delta \cap B \). This leads to an induced linear mapping from \( (24.10.13) \) into iself. More precisely, this induced linear mapping is nilpotent, because \( \text{ad}_x \) is nilpotent.

The collection of linear mappings from \( (24.10.13) \) into itself induced by \( \text{ad}_x \), \( x \in N \), is a Lie subalgebra of the algebra of all linear mappings from \( (24.10.13) \) into itself, because \( N \) is a Lie algebra over \( k \). Note that \( (24.10.13) \) is nontrivial, by \( (24.10.3) \). It follows that there is a nonzero element of \( (24.10.13) \) that is sent to 0 by the mapping induced by \( \text{ad}_x \) for every \( x \in N \), as in Section 9.9.

Equivalently, this means that there is a \( y \in B_\Delta \) such that \( y \notin B \) and \( \text{ad}_x(y) \in B_\Delta \cap B \) for every \( x \in N \). We also have that \( \text{ad}_x(y) \in N_\Delta \) for every \( x \in N \), because \( N_\Delta \) is an ideal in \( B_\Delta \). This implies that

\[
(24.10.14) \quad [y, x] = -\text{ad}_x(y) \in N
\]

for every \( x \in N \), so that \( y \in M \). It follows that \( (24.10.12) \) holds, because \( y \in B_\Delta \) and \( y \notin B \).

Similarly, let us verify that

\[
(24.10.15) \quad B_\Delta \cap B \neq B \cap M.
\]

The quotient \( B/(B_\Delta \cap B) \) is a nontrivial vector space over \( k \), on which \( \text{ad}_x \) induces a linear mapping for every \( x \in N \). Using the same type of argument as before, we get that there is a \( z \in B \) such that \( z \notin B_\Delta \) and \( \text{ad}_x(z) \in B_\Delta \cap B \) for every \( x \in N \).

If \( x \in N \), then \( [x, z] \in [B, B] \), because \( N \subseteq B \) and \( z \in B \). This means that

\[
(24.10.16) \quad [x, z] \in B_\Delta \cap [B, B] \subseteq N,
\]

using \( (24.10.10) \) in the second step. It follows that \( z \in M \). Thus \( z \) is an element of the right side of \( (24.10.15) \), and not the left.
24.10. THE FIRST SUBCASE

24.10.2 Some Borel subalgebras of $M$ and $A$

Note that $B_{\Delta} \cap M$, $B \cap M$ are solvable as Lie algebras over $k$, because $B_{\Delta}$, $B$ are solvable. Let $C_{\Delta}$, $C$ be Borel subalgebras of $M$ such that

\[(24.10.17) \quad B_{\Delta} \cap M \subseteq C_{\Delta}, \quad B \cap M \subseteq C.\]

Remember that $M$ is a proper Lie subalgebra of $A$, so that the dimension of $M$ is less than the dimension of $A$. Thus our induction hypothesis implies that there is an element of $\mathcal{E}(M)$ that maps $C$ onto $C_{\Delta}$.

Every element of $\mathcal{E}(M)$ can be obtained by restricting an element of $\mathcal{E}(A, M)$ to $M$. It follows that there is a $\sigma \in \mathcal{E}(A, M) \subseteq \mathcal{E}(A)$ such that

\[(24.10.18) \quad \sigma(C) = C_{\Delta}.\]

Let $B_1$ be a Borel subalgebra of $A$ that contains $C_{\Delta}$. Note that

\[(24.10.19) \quad B_{\Delta} \cap B \subseteq B_{\Delta} \cap M \subseteq B_{\Delta} \cap C_{\Delta} \subseteq B_{\Delta} \cap B_1,\]

by (24.10.11) and (24.10.17). We also have that

\[(24.10.20) \quad B_{\Delta} \cap B \neq B_{\Delta} \cap B_1,\]

by (24.10.12). This implies that the dimension of $B_{\Delta} \cap B$ is strictly less than the dimension of $B_{\Delta} \cap B_1$. This permits us to use our second induction hypothesis, to get an element $\tau$ of $\mathcal{E}(A)$ such that

\[(24.10.21) \quad \tau(B_1) = B_{\Delta}.\]

Combining (24.10.18) and (24.10.21), we obtain that

\[(24.10.22) \quad \tau(\sigma(C)) = \tau(C_{\Delta}) \subseteq \tau(B_1) = B_{\Delta}.\]

Observe that

\[(24.10.23) \quad \tau(\sigma(B_{\Delta} \cap B)) \subseteq \tau(\sigma(B \cap M)) \subseteq \tau(\sigma(C)) \cap \tau(\sigma(B)) \subseteq B_{\Delta} \cap \tau(\sigma(B)),\]

by (24.10.11), (24.10.17), and (24.10.22). We also have that the first inclusion is strict, by (24.10.15).

This shows that the dimension of $B_{\Delta} \cap B$ is strictly less than the dimension of $B_{\Delta} \cap \tau(\sigma(B))$. Of course, $\tau(\sigma(B))$ is a Borel subalgebra of $A$ too. Thus we may use our second induction hypothesis again, to get that there is an element of $\mathcal{E}(A)$ that maps $B_{\Delta}$ onto $\tau(\sigma(B))$. This means that there is an element of $\mathcal{E}(A)$ that maps $B_{\Delta}$ onto $B$, as desired.
24.11 The second subcase

We continue with the same notation and hypotheses as in Section 24.9. In this
section, we continue to suppose that \( B_{\Delta} \cap B \neq \{0\} \) as well, and now also that

\[
B_{\Delta} \cap B \text{ does not have any nonzero elements that are ad-nilpotent in } A.
\]

This corresponds to Case (ii) of (1) on p85 of [13].

Let \( B_0 \) be a Borel subalgebra of \( A \), and let \( x \in B_0 \) be given. Remember that there are unique \( x_1, x_2 \in A \) such that \( x = x_1 + x_2 \), \([x_1, x_2]_A = 0\), and \( \text{ad}_{x_1}, \text{ad}_{x_2} \) are the diagonalizable and nilpotent parts of \( \text{ad}_x \), as a linear mapping from \( A \) into itself, as in Section 14.3. We also have that

\[
\text{ad}_{x_1}(B_0), \text{ad}_{x_2}(B_0) \subseteq B_0,
\]

because \( \text{ad}_{x}(B_0) \subseteq B_0 \), as in Section 10.8. This means that \( x_1, x_2 \) are elements of the normalizer of \( B_0 \) in \( A \). It follows that

\[
x_1, x_2 \in B_0,
\]

because the normalizer of a Borel subalgebra of \( A \) is itself, as in Section 22.7.

If \( x \in B_{\Delta} \cap B \), then the ad-diagonalizable and ad-nilpotent parts of \( x \) are elements of \( B_{\Delta} \cap B \) too, by the remarks in the preceding paragraph. This implies that the ad-nilpotent part of \( x \) is 0, by hypothesis. This means that every element of \( B_{\Delta} \cap B \) is ad-diagonalizable in \( A \) under these conditions.

Of course,

\[
T = B_{\Delta} \cap B
\]

is a Lie subalgebra of \( A \), and in fact a toral subalgebra of \( A \), as in Section 17.1. In particular, \( T \) is commutative as a Lie algebra over \( k \), as before. It follows that \( \text{ad}_{x}, x \in T \), can be simultaneously diagonalized on \( A \).

Let \( N_{\Delta} \) be as in (24.10.4), and note that

\[
N_{\Delta} \cap T = \{0\},
\]

by hypothesis. Thus

\[
[B_{\Delta}, B_{\Delta}] \cap T = \{0\},
\]

by (24.10.5). Of course, this implies that

\[
[B_{\Delta}, T] \cap T = \{0\},
\]

because \( T \subseteq B_{\Delta} \). It follows that the normalizer of \( T \) in \( B_{\Delta} \) is the same as the centralizer of \( T \) in \( B_{\Delta} \), which is to say that

\[
N_{B_{\Delta}}(T) = C_{B_{\Delta}}(T).
\]
24.11. THE SECOND SUBCASE

24.11.1 Reducing to \( T \subseteq A_0 \)

Let \( C \) be a Cartan subalgebra of \( C_{B_\Delta}(T) \), as a Lie algebra over \( k \). Observe that \( T \) is contained in the normalizer \( N_{C_{B_\Delta}(T)}(C) \) of \( C \) in \( C_{B_\Delta}(T) \), by definition of \( C_{B_\Delta}(T) \). This implies that

\[(24.11.9) \quad T \subseteq C,\]

because \( N_{C_{B_\Delta}(T)}(C) = C \), by the definition of a Cartan subalgebra.

Of course, \( C_{B_\Delta}(T) \subseteq B_\Delta \), by definition. Let \( u \) be an element of the normalizer \( N_{B_\Delta}(C) \) of \( C \) in \( B_\Delta \), and let \( t \in T \) be given. Clearly

\[(24.11.10) \quad \text{ad}_t(u) = [t, u]_A \in C,\]

because \( t \in C \). This implies that

\[(24.11.11) \quad \text{ad}_t(\text{ad}_t(u)) = 0,\]

because \( C \subseteq C_{B_\Delta}(T) \). It follows that

\[(24.11.12) \quad \text{ad}_t(u) = 0,\]

because \( \text{ad}_t \) is diagonalizable on \( A \), as before.

This shows that \( u \in C_{B_\Delta}(T) \), so that

\[(24.11.13) \quad N_{B_\Delta}(C) \subseteq C_{B_\Delta}(T).\]

Using this, it is easy to see that

\[(24.11.14) \quad N_{B_\Delta}(C) = N_{C_{B_\Delta}(T)}(C),\]

by definition of the normalizer. This means that

\[(24.11.15) \quad N_{B_\Delta}(C) = C,\]

because \( C \) is a Cartan subalgebra of \( C_{B_\Delta}(T) \). Of course, \( C \) is nilpotent as a Lie algebra over \( k \), because it is a Cartan subalgebra of \( C_{B_\Delta}(T) \). It follows that \( C \) is a Cartan subalgebra of \( B_\Delta \) as well.

Remember that \( A_0 \) is a maximal toral subalgebra of \( A \), as in Section 24.9. This implies that \( A_0 \) is a Cartan subalgebra of \( A \), as in Section 18.10. We also have that \( A_0 \subseteq B_\Delta \), by the definition (24.9.9) of \( B_\Delta \). It follows that \( A_0 \) is a Cartan subalgebra of \( B_\Delta \), as in Section 18.11.

Because \( B_\Delta \) is solvable as a Lie algebra over \( k \), and \( A_0, C \) are Cartan subalgebras of \( B_\Delta \), there is an element of \( \mathcal{E}(B_\Delta) \) that maps \( C \) onto \( A_0 \), as in Section 24.7. This implies that there is an element of \( \mathcal{E}(A, B_\Delta) \subseteq \mathcal{E}(A) \) that maps \( C \) onto \( A_0 \), as in Section 24.3. In particular, this element of \( \mathcal{E}(A, B_\Delta) \) maps \( T \) into \( A_0 \). Using this, we can reduce to the case where

\[(24.11.16) \quad T \subseteq A_0.\]
24.11.2 The case where $T = A_0$

Suppose for the moment that $T = A_0$. Remember that $A_0 \neq \emptyset$, as in Section 24.9. We also have $A \neq A_0$, because $A$ is semisimple, and $A_0$ is commutative as a Lie algebra. This implies that $\Phi \neq \emptyset$, so that $B_\Delta \neq A_0$. Thus $A_0$ is not a Borel subalgebra of $A$, and $A_0 \neq B$.

If $T = A_0$, then $A_0 \subseteq B$. If $x \in A_0$, then it follows that $ad_x$ maps $B$ into itself. This implies that the restrictions of the $ad_x$'s, $x \in A_0$, to $B$ are simultaneously diagonalizable on $B$. Remember that for each $\alpha \in \Phi$, $A_\alpha$ has dimension one as a vector space over $k$. This means that $B$ corresponds, as a vector space over $k$, to the direct sum of $A_0$ and some collection of $A_\alpha$'s. In particular, there is at least one $\alpha_0 \in \Phi$ such that

$$A_{\alpha_0} \subseteq B,$$

because $B \neq A_0$. More precisely, $\alpha_0 \not\in \Phi^+$, because $B_\Delta \cap B = T \subseteq A_0$. Thus $\alpha_0 \in -\Phi^+$.

Let $\theta_{\alpha_0}$ be the Lie algebra automorphism of $A$ corresponding to $\alpha_0$ as in (23.8.1). Note that the Lie subalgebra of $A$ denoted $B$ in Section 23.8 corresponds to $A_0$ here. It is easy to see that $\theta_{\alpha_0} \in E(A)$, because $x_\alpha, y_\alpha$ are strongly ad-nilpotent in $A$, as mentioned in Section 23.9. We also have that

$$\theta_{\alpha_0}(A_0) = A_0$$

and

$$\theta_{\alpha_0}(A_{\alpha_0}) = A_{-\alpha_0}.$$ 

More precisely, (24.11.18) corresponds to (23.8.4), and can also be obtained from (23.8.6), while (24.11.19) can be obtained from (23.7.9) or (23.8.6).

It follows that $\theta_{\alpha_0}(B)$ is a Borel subalgebra of $A$ with

$$A_0 + A_{-\alpha_0} \subseteq \theta_{\alpha_0}(B).$$

Of course, $A_0 + A_{-\alpha_0} \subseteq B_\Delta$, because $-\alpha_0 \in \Phi^+$. Thus

$$A_0 + A_{-\alpha_0} \subseteq B_\Delta \cap \theta_{\alpha_0}(B).$$

In particular, the dimension of $B_\Delta \cap \theta_{\alpha_0}(B)$ is strictly larger than the dimension of $B_\Delta \cap B = T = A_0$, so that our second induction hypothesis implies that there is an element of $E(A)$ that maps $B_\Delta$ onto $\theta_{\alpha_0}(B)$. This implies that there is an element of $E(A)$ that maps $B_\Delta$ onto $B$, as desired.

24.11.3 The case where $T \neq A_0$, $B \subseteq C_A(T)$

Thus we suppose from now on in this section that

$$T \neq A_0.$$ 

Suppose for the moment that

$$B \subseteq C_A(T).$$
Remember that \( T = B_\Delta \cap B \neq \{0\} \), by hypothesis. This implies that

\[(24.11.24) \quad C_A(T) \neq A,\]

because otherwise \( T \) would be contained in the center of \( A \), which would contradict the semisimplicity of \( A \).

Observe that

\[(24.11.25) \quad A_0 \subseteq C_A(T),\]

because \( T \subseteq A_0 \) and \( A_0 \) is commutative as a Lie algebra. Let \( B_1 \) be a Borel subalgebra of \( C_A(T) \), as a Lie algebra over \( k \), that contains \( A_0 \). Of course, \( B \) may be considered as a Borel subalgebra of \( C_A(T) \) too, as in Section 18.11. Thus our first induction hypothesis implies that there is an element \( \sigma \) of \( E(A, C_A(T)) \) that maps \( B \) onto \( B_1 \). It follows that there is an element \( \sigma \) of \( E(A) \) that maps \( B_\Delta \) onto \( B_1 \), by (24.11.26).

\subsection*{24.11.4 The case where \( T \neq A_0, \ B \not\subseteq C_A(T) \)}

Suppose for the rest of the section that (24.11.23) does not hold. This means that there is a \( t_1 \in T \) such that \( \text{ad}_{t_1} \) is not identically zero on \( B \). Note that \( \text{ad}_{t_1} \) maps \( B \) into itself, because \( T \subseteq B \) and \( B \) is a Lie subalgebra of \( A \). We also have that \( \text{ad}_{t_1} \) is diagonalizable on \( B \), because it is diagonalizable on \( A \). It follows that there is an \( a_1 \in k \) and \( x_1 \in B \) such that \( a_1 \neq 0, x_1 \neq 0 \), and

\[(24.11.28) \quad \text{ad}_{t_1}(x_1) = a_1 x_1.\]

Remember that if \( \Psi_0 \subseteq \Phi \cup \{0\} \), then \( A(\Psi_0) \) is the linear subspace of \( A \) spanned by \( A_\alpha, \alpha \in \Psi_0 \), as in Subsection 24.9.2. Put

\[(24.11.29) \quad \Psi_1 = \{ \alpha \in \Phi : \alpha(t_1) = a_1 \}\]

and

\[(24.11.30) \quad \Psi_2 = \{ \alpha \in \Phi : \alpha(t_1) = n \cdot a_1 \text{ for some } n \in \mathbb{Z}_+ \},\]

so that \( \Psi_1 \subseteq \Psi_2 \). One can use (24.11.28) to check that

\[(24.11.31) \quad x_1 \in A(\Psi_1).\]
One can also verify that $A(\Psi_2), A(\Psi_2 \cup \{0\})$ are Lie subalgebras of $A$, with $A(\Psi_2)$ nilpotent, and $A(\Psi_2 \cup \{0\})$ solvable. More precisely, this can be obtained as in Section 22.5.

Let $B_2$ be a Borel subalgebra of $A$ that contains $A(\Psi_2 \cup \{0\})$. Thus

$$A_0 \subseteq B_\Delta \cap B_2,$$

which implies that the dimension of $B_\Delta \cap B = T$ is strictly less than the dimension of $B_\Delta \cap B_2$, because $T$ is a proper linear subspace of $A_0$. Using our second induction hypothesis, we get that there is an element $\sigma_2$ of $\mathcal{E}(A)$ such that

$$\sigma_2(B_\Delta) = B_2.$$

Observe that

$$x_1 \in B_2 \cap B,$$

using (24.11.31). Clearly $T \subseteq A_0 \subseteq B_2$, so that

$$T \subseteq B_2 \cap B.$$

Of course, $x_1$ is not in $T$, by (24.11.28). This implies that $B_\Delta \cap B = T$ is a proper linear subspace of $B_2 \cap B$. In particular, the dimension of $B_\Delta \cap B$ is strictly less than the dimension of $B_\Delta \cap B_2$.

It follows that the dimension of $B_\Delta \cap B$ is strictly less than the dimension of

$$\sigma_2^{-1}(B_2 \cap B) = B_\Delta \cap \sigma_2^{-1}(B).$$

Note that $\sigma_2^{-1}(B)$ is a Borel subalgebra of $A$, so that our second induction hypothesis implies that there is an element of $\mathcal{E}(A)$ that maps $B_\Delta$ onto $\sigma_2^{-1}(B)$. This implies that there is an element of $\mathcal{E}(A)$ that maps $B_\Delta$ onto $B$, as desired.

### 24.12 The second case

We continue with the same notation and hypotheses as in Section 24.9 again. The arguments in the previous two sections show that if $B_\Delta \cap B \neq \{0\}$, then there is an element of $\mathcal{E}(A)$ that maps $B_\Delta$ onto $B$. More precisely, this works for any choice of maximal toral subalgebra $A_0$ of $A$, and base $\Delta$ for the corresponding root system $\Phi$, as in Subsection 24.9.3.

In this section, we suppose that

$$B_\Delta \cap B = \{0\}.$$

This corresponds to (2) on p86 of [13].

Remember that $A$ corresponds to the direct sum of $A_\alpha$, $\alpha \in \Phi \cup \{0\}$, as a vector space over $k$, as in Subsection 24.9.2. In particular, the dimension of $A$, as a vector space over $k$, is given by

$$\dim A = \dim A_0 + \# \Phi.$$
where \(\#\Phi\) is the number of elements of \(\Phi\), because \(A_\alpha\) has dimension one as a vector space over \(k\). Similarly,

\[
(24.12.3) \quad \dim B_\Delta = \dim A(\Phi^+ \cup \{0\}) = \dim A_0 + \#\Phi^+.
\]

Note that

\[
(24.12.4) \quad \#\Phi^+ = \#\Phi/2,
\]

because \(\Phi\) is the union of \(\Phi^+\) and \(\Phi^-\), which are disjoint. Thus

\[
(24.12.5) \quad \dim B_\Delta = \dim A_0 + \#\Phi/2.
\]

It follows that

\[
(24.12.6) \quad (\dim A)/2 < \dim B_\Delta,
\]

because \(A_0 \neq \{0\}\). We also have that

\[
(24.12.7) \quad \dim B_\Delta + \dim B \leq \dim A,
\]

by (24.12.1). This means that

\[
(24.12.8) \quad \dim B < (\dim A)/2.
\]

Suppose for the moment that every element of \(B\) is ad-nilpotent in \(A\). Of course, this implies that every element of \(B\) is ad-nilpotent in \(B\). It follows that \(B\) is nilpotent as a Lie algebra over \(k\), as in Section 9.10.

Remember that the normalizer of \(B\) in \(A\) is equal to \(B\), because \(B\) is a Borel subalgebra of \(A\), as in Section 22.7. This means that \(B\) is a Cartan subalgebra of \(A\).

Under these conditions, we get that \(B\) is a toral subalgebra of \(A\), as in Section 18.10. Thus the elements of \(B\) are ad-diagonalizable in \(A\). This implies that \(B = \{0\}\), because the elements of \(B\) are also supposed to be ad-nilpotent in \(A\). This contradicts the hypothesis that \(B\) be a Borel subalgebra of \(A\), since \(A \neq \{0\}\).

Thus we may suppose now that \(B\) has an element that is not ad-nilpotent in \(A\). Remember that \(B\) contains the ad-diagonalizable and ad-nilpotent parts of its elements, as in (24.11.3). It follows that \(B\) has a nonzero element \(x_0\) that is ad-diagonalizable in \(A\).

Let \(T_0\) be a maximal toral subalgebra of \(A\) that contains \(x_0\). Using any base for the root system corresponding to \(T_0\), we get a standard Borel subalgebra \(B_0\) of \(A\) with respect to \(T_0\), as in Subsection 24.9.2. Note that \(T_0 \subseteq B_0\), so that

\[
(24.12.9) \quad B_0 \cap B \neq \{0\}.
\]

This implies that there is an element of \(E(A)\) that maps \(B_0\) onto \(B\), as mentioned at the beginning of the section. In particular,

\[
(24.12.10) \quad \dim B = \dim B_0.
\]

However, \(\dim B_0 > (\dim A)/2\), as in (24.12.6). This contradicts (24.12.8).

This shows that (24.12.1) is not possible. It follows that there is an element of \(E(A)\) that maps \(B_\Delta\) onto \(B\). This completes the proof of the theorem stated at the beginning of Section 24.9.
24.13 Some consequences

Let $k$ be an algebraically closed field of characteristic 0 again, and let $(A, [\cdot, \cdot]_A)$ be a finite-dimensional Lie algebra over $k$. The rank of $A$ as a Lie algebra over $k$ may be defined as the dimension of any Cartan subalgebra of $A$, as on p86 of [13]. The corollary stated at the beginning of Section 24.9 implies that

(24.13.1) the rank of $A$ does not depend on the choice of the Cartan subalgebra.

The rank of a finite-dimensional complex Lie algebra is defined another way in Definition 2 on p11 of [23], and Corollary 1 on p13 of [23] states that these two definitions are equivalent.

Suppose now that $A \neq \{0\}$ is semisimple, as a Lie algebra over $k$. Let $A_0$ be a maximal toral subalgebra of $A$, and let $\Phi$ be the corresponding root system, as in Subsection 24.9.2. Also let $\Delta$ be a base for $\Phi$, and $B_\Delta$ be the standard Borel subalgebra of $A$ associated to $A_0$ and $\Delta$, as in Subsection 24.9.2.

Suppose that $N$ is a Lie subalgebra of $A$ such that every element of $N$ is $\text{ad}$-nilpotent as an element of $A$. In particular, every element of $N$ is $\text{ad}$-nilpotent as an element of $N$, so that $N$ is nilpotent as a Lie algebra over $k$, as in Section 9.10. Of course, this implies that $N$ is solvable as a Lie algebra over $k$, so that there is a Borel subalgebra $B$ of $A$ that contains $N$.

The theorem stated at the beginning of Section 24.9 implies that there is an element $\sigma$ of $\mathcal{E}(A)$ such that $\sigma(B) = B_\Delta$. In particular, $\sigma(N) \subseteq B_\Delta$. In fact,

(24.13.2) $\sigma(N) \subseteq N_\Delta,$

where $N_\Delta$ is as in (24.10.4). This follows from (24.10.6), because the elements of $\sigma(N)$ are $\text{ad}$-nilpotent in $A$. This corresponds to Exercise 2 on p87 of [13], and the corollary to Theorem 5 on p48 of [23].

Note that the standard Borel subalgebras of $A$ associated to $A_0$ and any base for $\Phi$ contain $A_0$, by construction. Conversely, let $B$ be any Borel subalgebra of $A$ such that

(24.13.3) $A_0 \subseteq B$.

Under these conditions,

(24.13.4) $B$ is a standard Borel subalgebra of $A$ with respect to $A_0$

and some base for $\Phi$,

as mentioned on p86 of [13].

To see this, let $\Delta$ be a base for $\Phi$ again, and let $B_\Delta$ be the corresponding standard Borel subalgebra of $A$, as before. Remember that there is an element $\sigma$ of $\mathcal{E}(A)$ such that $\sigma(B_\Delta) = B$, by the theorem stated at the beginning of Section 24.9. In particular,

(24.13.5) $\sigma(A_0) \subseteq \sigma(B_\Delta) \subseteq B$. 

Remember that $A_0$ is a Cartan subalgebra of $A$, as in Section 18.10. Of course, this implies that $\sigma(A_0)$ is a Cartan subalgebra of $A$ too. It follows that $A_0$ and $\sigma(A_0)$ may be considered as Cartan subalgebras of $B$, as in Section 18.11.

Thus there is an element of $\mathcal{E}(B)$ that maps $\sigma(A_0)$ onto $A_0$, as in Section 24.7. This implies that there is an element $\tau$ of $\mathcal{E}(A, B) \subseteq \mathcal{E}(A)$ such that

\begin{equation}
\tau(\sigma(A_0)) = A_0,
\end{equation}

as in Section 24.3. Note that

\begin{equation}
\tau(\sigma(B_\Delta)) = \tau(B) = B.
\end{equation}

Put $\sigma_1 = \tau \circ \sigma \in \mathcal{E}(A)$, so that $\sigma_1(A_0) = A_0$ and $\sigma_1(B_\Delta) = B$. Let $\sigma_0$ be the restriction of $\sigma_1$ to $A_0$, and let $\sigma'_0$ be the corresponding dual linear mapping from the dual $A'_0$ of $A_0$ into itself. If $\alpha \in A'_0$ and $A_\alpha$ is as in Subsection 24.9.2, then

\begin{equation}
\sigma_1(A_{\sigma'_0(\alpha)}) = A_\alpha
\end{equation}

for every $\alpha \in A'_0$, as in Section 23.6. In particular,

\begin{equation}
\sigma'_0(\Phi) = \Phi,
\end{equation}

as before.

Let $E_Q$ be the linear subspace of $A'_0$, as a vector space over $Q$, spanned by $\Phi$, as in Subsection 24.9.2. Thus

\begin{equation}
\sigma'_0(E_Q) = E_Q,
\end{equation}

by (24.13.9), as in Section 23.6. This leads to a one-to-one linear mapping from the corresponding real vector space $E_R$ onto itself, which is an automorphism of $\Phi$ as a root system in $E_R$, as before. Using this, one can check that

\begin{equation}
\tilde{\Delta} = (\sigma'_0)^{-1}(\Delta)
\end{equation}

is a base for $\Phi$ as a root system in $E_R$ too.

Let $B_\tilde{\Delta}$ be the standard Borel subalgebra of $A$ associated to $A_0$ and $\tilde{\Delta}$ as in Subsection 24.9.2. One can verify that

\begin{equation}
\sigma_1(B_\Delta) = B_\tilde{\Delta},
\end{equation}

using (24.13.8). This means that

\begin{equation}
B = B_\tilde{\Delta},
\end{equation}

as desired, because $\sigma_1(B_\Delta) = B$, by construction.
24.14 The case of $sl_2(k)$

Exercise 4 on p87 of [13] asks one to consider simplifications of the proof of the theorem stated at the beginning of Section 24.9, in the case where the Lie algebra is $sl_2(k)$. Let us begin with a preliminary remark.

Let $k$ be an algebraically closed field, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$ of finite dimension at least 2. Suppose that $x$ is a nonzero element of $A$, so that the quotient

\[(24.14.1) \quad A/\{tx : t \in k\}\]

may be considered as a vector space over $k$ of positive finite dimension. Note that $ad_x$ induces a linear mapping from (24.14.1) into itself. This mapping has a nonzero eigenvector in (24.14.1), because $k$ is algebraically closed.

Equivalently, there is an element $y$ of $A$ such that $x$ and $y$ are linearly independent in $A$, and $ad_x(y)$ can be expressed as a linear combination of $x$ and $y$. This means that the linear span of $x$ and $y$ is a 2-dimensional Lie subalgebra of $A$, which is solvable as a Lie algebra over $k$. It follows that

\[(24.14.2) \quad \text{every Borel subalgebra of } A \text{ has dimension at least two}\]

under these conditions.

Now let $k$ be an algebraically closed field of characteristic 0, as in Section 24.9. Note that the Borel subalgebras of $sl_2(k)$ are proper subalgebras of $sl_2(k)$, because $sl_2(k)$ is not solvable. This implies that every Borel subalgebra of $sl_2(k)$ has dimension two as a vector space over $k$, by (24.14.2). Conversely, every two-dimensional Lie subalgebra of $sl_2(k)$ is solvable, and thus a Borel subalgebra.

Let $x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $y = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, and $h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ be the usual basis for $sl_2(k)$, as a vector space over $k$. Thus $[x,y] = h$, $[h,x] = 2 \cdot x$, and $[h,y] = -2 \cdot y$, as before. The linear span of $h$ is a maximal toral subalgebra of $sl_2(k)$, for which $\Phi$ corresponds to the linear functionals defined by multiplication by 2 and $-2$. The linear functional defined by multiplication by 2 may be considered as a base for $\Phi$ as a root system, for which the corresponding standard Borel subalgebra $B_\Delta$ of $sl_2(k)$ is spanned by $h$ and $x$.

Let $B$ be any Borel subalgebra of $sl_2(k)$, so that $B$ has dimension two as a vector space over $k$, as before. In particular, $B_\Delta \cap B \neq \{0\}$ in this case. We would like to show that there is an element of $\mathcal{L}(sl_2(k))$ that maps $B_\Delta$ onto $B$. Of course, this is trivial when $B = B_\Delta$, and so we may suppose that $B \neq B_\Delta$.

Suppose for the moment that $x \in B$. If $B \subseteq B_\Delta$, then $B = B_\Delta$, and so we may suppose that $B \nsubseteq B_\Delta$. Thus

\[(24.14.3) \quad ax + by + ch \in B\]

for some $a, b, c \in k$, with $b \neq 0$. This implies that

\[(24.14.4) \quad [x, ax + by + ch] = bh - 2cx \in B,\]

because $x \in B$ and $B$ is a Lie subalgebra of $sl_2(k)$. It follows that $h \in B$, because $x \in B$. This means that $y \in B$, by (24.14.3), which is to say that $B = sl_2(k)$. This is a contradiction, because $sl_2(k)$ is not solvable.
Thus we may suppose that \( x \notin B \). It follows that
\[
(24.14.5) \quad ax + h \in B
\]
for some \( a \in k \), because \( B_\Delta \cap B \neq \{0\} \). One can use the exponential of a multiple of \( \text{ad}_x \) to reduce to the case where \( h \in B \). Using this, one can check that either \( B = B_\Delta \), or \( B \) is spanned by \( y \) and \( h \). If \( B \) is spanned by \( y \) and \( h \), then \( B = \sigma(B_\Delta) \), where \( \sigma \in \mathcal{E}(\mathfrak{sl}_2(k)) \) is as in Section 15.7.

### 24.15 Automorphisms and semisimplicity

Let \( k \) be an algebraically closed field of characteristic 0, and let \( (A, [[\cdot, \cdot], A]) \) be a semisimple Lie algebra over \( k \) of positive finite dimension. Also let \( A_0 \) be a maximal toral subalgebra of \( A \), and let \( \Phi \) be the corresponding root system, as in Subsection 24.9.2. Choose a base \( \Delta \) for \( \Phi \), and let \( B_\Delta \) be the standard Borel subalgebra of \( A \) associated to \( A_0 \) and \( \Delta \), as before.

Suppose that \( \tau \) is an automorphism of \( A \) as a Lie algebra over \( k \). Thus \( \tau(B) \) is another Borel subalgebra of \( A \). The theorem stated at the beginning of Section 24.9 implies that there is an element \( \sigma_1 \) of \( \mathcal{E}(A) \) such that
\[
(24.15.1) \quad \sigma_1(\tau(B_\Delta)) = B_\Delta.
\]
Remember that \( A_0 \) is a Cartan subalgebra of \( A \), as in Section 18.10. This implies that \( \sigma_1(\tau(A_0)) \) is a Cartan subalgebra of \( A \) as well. It follows that \( A_0 \) and \( \sigma_1(\tau(A_0)) \) are Cartan subalgebras of \( B_\Delta \) too, as a Lie algebra over \( k \), as in Section 18.11. Thus there is an element of \( \mathcal{E}(B_\Delta) \) that maps \( \sigma_1(\tau(A_0)) \) onto \( A_0 \), as in Section 24.7. This means that there is an element \( \sigma_2 \) of \( \mathcal{E}(A, B_\Delta) \subseteq \mathcal{E}(A) \) with
\[
(24.15.2) \quad \sigma_2(\sigma_1(\tau(A_0))) = A_0,
\]
as in Section 24.3.

Put
\[
(24.15.3) \quad \tau_2 = \sigma_2 \circ \sigma_1 \circ \tau,
\]
which is an automorphism of \( A \) that maps \( A_0 \) onto itself. Let \( \tau_{2,0} \) be the restriction of \( \tau_2 \) to \( A_0 \), and let \( \tau'_{2,0} \) be the corresponding dual linear mapping from \( A_0' \) into itself. Thus
\[
(24.15.4) \quad \tau_2(A_{\tau'_{2,0}(\alpha)}) = A_\alpha
\]
for every \( \alpha \in A_0' \), as in Section 23.6, where \( A_\alpha \) is as in Subsection 24.9.2. This implies that
\[
(24.15.5) \quad \tau'_{2,0}(\Phi) = \Phi,
\]
as before.

It follows that
\[
(24.15.6) \quad \tau'_{2,0}(E_Q) = E_Q,
\]
as in Section 23.6, where \( E_Q \) is the linear subspace of \( A_0' \), as a vector space over \( Q \), spanned by \( \Phi \). As before, this leads to a one-to-one linear mapping from the
corresponding real vector space $E_{\mathbb{R}}$ onto itself, which is an automorphism of $\Phi$ as a root system in $E_{\mathbb{R}}$.

Note that
\[(24.15.7) \quad \tau_2(B_\Delta) = B_\Delta,\]
by (24.15.1), and because $\sigma_2 \in \mathcal{E}(A, B_\Delta)$. Let $\Phi^+$ be the set of positive roots with respect to $\Delta$, as in Subsection 24.9.2. One can check that
\[(24.15.8) \quad \tau'_{2,0}(\Phi^+) = \Phi^+,\]
using (24.15.4) and (24.15.7).

Using (24.15.8), one can verify that
\[(24.15.9) \quad \tau'_{2,0}(\Delta) = \Delta.\]

More precisely, $\Delta$ consists of exactly the elements of $\Phi^+$ that cannot be expressed as the sum of at least two other elements of $\Phi^+$.

We shall see in Section 28.6 that every automorphism of $\Phi$ as a root system that sends $\Delta$ onto itself corresponds to an automorphism of $A$ that maps $A_0$ onto itself. This also uses the remarks in Section 28.7.

The automorphisms of $A$ that are equal to the identity mapping on $A_0$ are called diagonal automorphisms, and they will be discussed further in Sections 28.6 and 29.7.

This corresponds to some remarks on p87 of [13].
Part IV

Some related constructions
Chapter 25

Universal enveloping algebras

25.1 Polynomial algebras

Let $I$ be a nonempty set, and let $\mathcal{M}_I$ be the set of functions $\alpha$ on $I$ with values in the set $\mathbb{Z}_+ \cup \{0\}$ of nonnegative integers such that $\alpha(j) = 0$ for all but finitely many $j \in I$. The elements of $\mathcal{M}_I$ are considered as \textit{multi-indices} with respect to $I$, which reduces to the usual notion when $I = \{1, \ldots, n\}$ for some positive integer $n$. If $\alpha \in \mathcal{M}_I$, then we put

\begin{equation}
|\alpha| = \sum_{j \in I} \alpha(j),
\end{equation}

which reduces to a finite sum. If $\beta \in \mathcal{M}_I$ too, then $\alpha + \beta \in \mathcal{M}_I$, and

\begin{equation}
|\alpha + \beta| = |\alpha| + |\beta|.
\end{equation}

Let $X_j$, $j \in I$, be a commuting family of indeterminates. If $\alpha \in \mathcal{M}_I$, then $X^\alpha$ is considered as a formal monomial in the $X_j$'s of degree $|\alpha|$. This may be identified with the formal product

\begin{equation}
X_{j_1}^{\alpha(j_1)} \cdots X_{j_n}^{\alpha(j_n)}
\end{equation}

when $j_1, \ldots, j_n$ are finitely many distinct elements of $I$, and $\alpha(j) = 0$ for every $j \in I \setminus \{j_1, \ldots, j_n\}$. If $\beta \in \mathcal{M}_I$ as well, then the corresponding product of formal monomials can be defined as usual by

\begin{equation}
X^\alpha \cdot X^\beta = X^{\alpha + \beta}.
\end{equation}

Let $k$ be a commutative ring with a multiplicative identity element. A \textit{formal polynomial} in the $X_j$'s, $j \in I$, with coefficients in $k$ can be expressed as

\begin{equation}
f(X) = \sum_{\alpha \in \mathcal{M}_I} f_\alpha X^\alpha,
\end{equation}

526
25.1. POLYNOMIAL ALGEBRAS

where \( f_\alpha \in k \) for every \( \alpha \in M_I \), and \( f_\alpha = 0 \) for all but finitely many \( \alpha \in M_I \). The space \( k[[X_j : j \in I]] \) of these formal polynomials can be defined as the space \( c_{00}(M_I, k) \) of all \( k \)-valued functions on \( M_I \) with finite support, which is to say that are equal to 0 at all but finitely many elements of \( M_I \). This is a submodule of the space \( c(M_I, k) \) of all \( k \)-valued functions on \( M_I \), as a module over \( k \) with respect to pointwise addition and scalar multiplication of functions.

Of course, this corresponds to termwise addition and scalar multiplication of formal polynomials as in (25.1.5).

Let \( g(X) = \sum_{\beta \in M_I} g_\beta X^\beta \)

be another formal polynomial in the \( X_j \)'s, \( j \in I \), with coefficients in \( k \). If \( \gamma \in M_I \), then put

\[
(25.1.7) \quad h_\gamma = \sum_{\alpha + \beta = \gamma} f_\alpha g_\beta,
\]

where more precisely the sum is taken over all \( \alpha, \beta \in M_I \) such that \( \alpha + \beta = \gamma \). It is easy to see that there are only finitely many such \( \alpha, \beta \in M_I \), so that the right side of (25.1.7) is a finite sum in \( k \). We also have that (25.1.7) is equal to 0 for all but finitely many \( \gamma \in M_I \), because of the analogous conditions for \( f_\alpha \) and \( g_\beta \). Thus

\[
(25.1.8) \quad h(X) = \sum_{\gamma \in M_I} h_\gamma X^\gamma
\]

is a formal polynomial in the \( X_j \)'s, \( j \in I \), with coefficients in \( k \), and we put \( f(X) g(X) = h(X) \). One can check that \( k[[X_j : j \in I]] \) is a commutative associative algebra over \( k \) with respect to this definition of multiplication. Let us identify elements of \( k \) with the formal polynomials whose coefficients of \( X^\alpha \) are equal to 0 when \( \alpha \) is a nonzero element of \( M_I \), and are equal to the given element of \( k \) when \( \alpha = 0 \). This is a subalgebra of \( k[[X_j : j \in I]] \), and 1 \( \in k \) corresponds to the multiplicative identity element in \( k[[X_j : j \in I]] \).

Let \( B \) be an associative algebra over \( k \) with a multiplicative identity element \( e_B \), and let \( b_j \) be an element of \( B \) for each \( j \in I \). The Cartesian power \( B^I \) is the same as the Cartesian product of the family of copies of \( B \) indexed by \( I \), and we let \( b \) be the element of \( B^I \) whose \( j \)th coordinate is equal to \( b_j \) for each \( j \in I \). Suppose that the \( b_j \)'s commute in \( B \), which is to say that

\[
(25.1.9) \quad b_j b_l = b_l b_j
\]

for every \( j, l \in I \). If \( \alpha \in M_I \), and \( j_1, \ldots, j_n \) are finitely many distinct elements of \( I \) such that \( \alpha(j) = 0 \) when \( j \in I \setminus \{j_1, \ldots, j_n\} \), then put

\[
(25.1.10) \quad b^\alpha = b_{j_1}^{\alpha(j_1)} \cdots b_{j_n}^{\alpha(j_n)}.
\]

As usual, \( b_j^{\alpha(j)} \) is interpreted as being \( e_B \) when \( \alpha(j) = 0 \), so that the right side of (25.1.10) does not depend on the choice of \( j_1, \ldots, j_n \). If \( \beta \in M_I \) too, then

\[
(25.1.11) \quad b^{\alpha + \beta} = b^\alpha b^\beta,
\]
because of (25.1.9). If \( f(X) \) is as in (25.1.5), then put
\[
(25.1.12) \quad f(b) = \sum_{\alpha \in \mathcal{M}_f} f_{\alpha} b^\alpha,
\]
which reduces to a finite sum in \( B \). It is easy to see that \( f(X) \mapsto f(b) \) defines
an algebra homomorphism from \( k[[X_j : j \in J]] \) into \( B \).

25.2 Tensor products of associative algebras

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( B_1, B_2 \) be
associative algebras over \( k \). Thus \( B_1, B_2 \) are modules over \( k \) in particular,
and we take \( B = B_1 \otimes B_2 \), as a module over \( k \). If \( b_1, b'_1 \in B_1 \) and \( b_2, b'_2 \in B_2 \), then
\[
(25.2.1) \quad (b_1 b'_1) \otimes (b_2 b'_2)
\]
is an element of \( B \), and this defines a multilinear mapping from \( B_1 \times B_2 \times B_1 \times B_2 \)
into \( B \). This multilinear mapping corresponds to a unique linear mapping from
\( B_1 \otimes B_2 \otimes B_1 \otimes B_2 \) into \( B \), as usual. Remember that \( B_1 \otimes B_2 \otimes B_1 \otimes B_2 \)
can be identified with \( B \otimes B \), as in Section 7.15. This leads to a bilinear
mapping from \( B \times B \) into \( B \), by composing the standard bilinear mapping from
\( B \times B \) into \( B \otimes B \) with the previous linear mapping from \( B \otimes B \) into \( B \). This
bilinear mapping is used to define multiplication on \( B \). By construction,
\[
(25.2.2) \quad (b_1 \otimes b_2) (b'_1 \otimes b'_2) = (b_1 b'_1) \otimes (b_2 b'_2)
\]
in \( B \) for every \( b_1, b'_1 \in B_1 \) and \( b_2, b'_2 \in B_2 \). One can check that multiplication
on \( B \) is associative, because multiplication is associative on \( B_1 \) and \( B_2 \), by
hypothesis. If multiplication is commutative on \( B_1 \) and \( B_2 \), then multiplication
on \( B \) is commutative as well.

Suppose for the moment that \( B_1 \) and \( B_2 \) have multiplicative identity elements \( e_1 \) and \( e_2 \), respectively. In this case, \( e_1 \otimes e_2 \) is the multiplicative identity
element in \( B \). Observe that
\[
(25.2.3) \quad b_1 \mapsto b_1 \otimes e_2
\]
and
\[
(25.2.4) \quad b_2 \mapsto e_1 \otimes b_2
\]
are algebra homomorphisms from \( B_1 \) and \( B_2 \) into \( B \). We also have that
\[
(25.2.5) \quad (b_1 \otimes e_2) (e_1 \otimes b_2) = (e_1 \otimes b_2) (b_1 \otimes e_2) = b_1 \otimes b_2
\]
for every \( b_1 \in B_1 \) and \( b_2 \in B_2 \).

Let \( C \) be another associative algebra over \( k \), and let \( \phi_1, \phi_2 \) be algebra homomorphisms from \( B_1, B_2 \) into \( C \), respectively. Suppose that
\[
(25.2.6) \quad \phi_1(b_1) \phi_2(b_2) = \phi_2(b_2) \phi_1(b_1)
\]
for every $b_1 \in B_1$ and $b_2 \in B_2$. Of course, $(b_1, b_2) \mapsto \phi_1(b_1) \phi_2(b_2)$ is a bilinear mapping from $B_1 \times B_2$ into $C$. This leads to a unique linear mapping $\phi$ from $B$ into $C$, with

\[(25.2.7)\quad \phi(b_1 \otimes b_2) = \phi_1(b_1) \phi_2(b_2)\]

for every $b_1 \in B_1$ and $b_2 \in B_2$. One can verify that $\phi$ is an algebra homomorphism from $B$ into $C$ under these conditions.

### 25.3 Tensor algebras

Let $k$ be a commutative ring with a multiplicative identity element, and let $V$ be a module over $k$. If $n$ is a positive integer, then let $V^n$ be the space of $n$-tuples of elements of $V$, which is the same as the Cartesian product of $n$ copies of $V$. Remember that the $n$th tensor power $T^n V$ of $V$ is defined as a module over $k$ as in Section 7.15, as the tensor product of $n$ copies of $V$. This is interpreted as being equal to $V$ when $n = 1$, as before. In particular, the mapping from $(v_1, \ldots, v_n) \in V^n$ to $v_1 \otimes \cdots \otimes v_n \in T^n V$ is multilinear over $k$.

Let $n_1$ and $n_2$ be positive integers, and remember that there is a natural isomorphism between $(T^{n_1} V) \otimes (T^{n_2} V)$ and $T^{n_1 + n_2} V$, as modules over $k$, as in Section 7.15. There is also a natural bilinear mapping from $T^{n_1} V \times T^{n_2} V$ into $(T^{n_1} V) \otimes (T^{n_2} V)$, by the definition of the tensor product. This leads to a natural bilinear mapping from $(T^{n_1} V) \times (T^{n_2} V)$ into $T^{n_1 + n_2} V$. If $a_1 \in T^{n_1} V$ and $a_2 \in T^{n_2} V$, then let $a_1 a_2$ be the image of $(a_1, a_2)$ in $T^{n_1 + n_2} V$ under this mapping. If $a_1 = v_1 \otimes \cdots \otimes v_{n_1}$ and $a_2 = w_1 \otimes \cdots \otimes w_{n_2}$ for some $v_1, \ldots, v_{n_1}$ and $w_1, \ldots, w_{n_2}$ in $V$, then

\[(25.3.1)\quad (v_1 \otimes \cdots \otimes v_{n_1}) (w_1 \otimes \cdots \otimes w_{n_2}) = v_1 \otimes \cdots \otimes v_{n_1} \otimes w_1 \otimes \cdots \otimes w_{n_2}\]

in $T^{n_1 + n_2} V$, by construction.

It is customary to define $T^n V$ when $n = 0$ to be $k$, as a module over itself. If $W$ is any module over $k$, then scalar multiplication on $W$ can be used to define bilinear mappings from $k \times W$ and from $W \times k$ into $W$. This leads to standard isomorphisms from $k \otimes W$ and from $W \otimes k$ onto $W$, as modules over $k$. In particular, we have a natural bilinear mapping from $T^{n_1} V \times T^{n_2} V$ into $T^{n_1 + n_2} V$ when $n_1 = 0$ or $n_2 = 0$ as well. If $a_1 \in T^{n_1} V$, $a_2 \in T^{n_2} V$, and $a_3 \in T^{n_3} V$ for some nonnegative integers $n_1, n_2, n_3$, then one can check that

\[(25.3.2)\quad (a_1 a_2) a_3 = a_1 (a_2 a_3)\]

in $T^{n_1 + n_2 + n_3} V$.

The tensor algebra on $V$ is denoted $TV$, and defined initially as a module over $k$ to be the direct sum of $T^n V$ over all nonnegative integers $n$,

\[(25.3.3)\quad TV = \bigoplus_{n=0} T^n V.\]

There is a natural bilinear mapping from $(TV) \times (TV)$ into $TV$, which is defined using the natural bilinear mapping from $(T^{n_1} V) \times (T^{n_2} V)$ into $T^{n_1 + n_2} V$ for
n_1, n_2 \geq 0$ mentioned in the previous two paragraphs. If $a, b \in TV$, then let $ab$ be the image of $(a, b)$ in $TV$ under this mapping. It is easy to see the multiplication on $TV$ is associative, using (25.3.2). We may consider $k = T^0V$ as a subalgebra of $TV$, and the multiplicative identity element $1$ in $k$ is the multiplicative identity element in $TV$ too.

Let $A$ be an associative algebra over $k$ with a multiplicative identity element $e$, and let $\phi$ be a homomorphism from $V$ into $A$, as modules over $k$. If $n$ is a positive integer, then there is a natural multilinear mapping from $V^n$ into $A$, which sends $(v_1, \ldots, v_n) \in V^n$ to $\phi(v_1) \cdots \phi(v_n)$. This leads to a unique linear mapping $\psi_n$ from $T^nV$ into $A$, with

\[
\psi_n(v_1 \otimes \cdots \otimes v_n) = \phi(v_1) \cdots \phi(v_n)
\]

for every $(v_1, \ldots, v_n) \in V^n$. Put $\psi_0(t) = te$ for every $t \in k = T^0V$. If $a_1 \in T^{n_1}V$ and $a_2 \in T^{n_2}V$ for some $n_1, n_2 \geq 0$, then one can check that

\[
\psi_{n_1+n_2}(a_1 a_2) = \psi_{n_1}(a_1) \psi_{n_2}(a_2).
\]

Let $\psi$ be the mapping from $TV$ into $A$ which corresponds to $\psi_n$ on $T^nV$ for every $n \geq 0$. This defines an algebra homomorphism from $TV$ into $A$, because of (25.3.5). More precisely, this is the unique algebra homomorphism from $TV$ into $A$ that is equal to $\phi$ on $V = T^1V$ and sends $1 \in k = T^0V$ to $e$. Note that every algebra homomorphism $\psi$ from $TV$ into $A$ with $\psi(1) = e$ occurs in this way, where $\phi$ corresponds to $\psi$ on $V = T^1V$. This corresponds to some of the remarks on p11 of [24], and on p89 of [13].

### 25.4 Defining universal enveloping algebras

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. A **universal enveloping algebra** of $A$ is an associative algebra $UA$ over $k$ with a multiplicative identity element $e = e_{UA}$ and a mapping $i = i_{UA}$ from $A$ into $UA$ with the following two properties. First, $i$ is a Lie algebra homomorphism from $A$ into $UA$, considered as a Lie algebra over $k$ with respect to the corresponding commutator bracket. This means that $i$ is a homomorphism from $A$ into $UA$ as modules over $k$ such that

\[
i([x, y]_A) = i(x) i(y) - i(y) i(x)
\]

for every $x, y \in A$.

Second, let $B$ be an associative algebra over $k$ with a multiplicative identity element $e_B$, and let $\phi$ be a Lie algebra homomorphism from $A$ into $B$, as a Lie algebra over $k$ with respect to the corresponding commutator bracket. As before, this means that $\phi$ is a homomorphism from $A$ into $B$, as modules over $k$, such that

\[
\phi([x, y]_A) = \phi(x) \phi(y) - \phi(y) \phi(x)
\]
for every \(x, y \in B\). Under these conditions, there should be a unique algebra homomorphism \(\psi\) from \(UA\) into \(B\) such that

\[
\psi(e_{UA}) = e_B
\]

and

\[
\psi \circ i = \phi
\]

on \(A\). This corresponds to Definition 1.1 on p11 of [24], and to the definition on p90 of [13].

In particular, we can apply the second condition to \(B = UA\) and \(\phi = i\). This implies that there is a unique algebra homomorphism \(\tilde{\psi}\) from \(UA\) into itself such that

\[
\tilde{\psi}(e_{UA}) = e_{UA}, \quad \psi \circ i_{UA} = i_{UA'}, \quad \tilde{\psi} \circ i_{UA} = i_{UA}
\]

and similarly \(\tilde{\psi} \circ i_{UA} = i_{UA}\). Of course,

\[
(\tilde{\psi} \circ \psi) \circ i_{UA} = \tilde{\psi} \circ (\psi \circ i_{UA}) = \tilde{\psi} \circ i_{UA} = i_{UA},
\]

and similarly \(\tilde{\psi} \circ \psi) \circ i_{UA} = i_{UA}\). Clearly \(\tilde{\psi} \circ \psi\) is an algebra homomorphism from \(UA\) into itself, and \(\psi \circ \tilde{\psi}\) is an algebra homomorphism from \(\tilde{U}A\) into itself. It follows that \(\tilde{\psi} \circ \psi\) is the identity mapping on \(UA\), and that \(\psi \circ \tilde{\psi}\) is the identity mapping on \(\tilde{U}A\), as in the preceding paragraph. This corresponds to some remarks on p11 of [24], and on p90f of [13].

### 25.5 Constructing universal enveloping algebras

Let \(k\) be a commutative ring with a multiplicative identity element, and let \((A, \cdot, [\cdot, \cdot])_A\) be a Lie algebra over \(k\). In particular, \(A\) is a module over \(k\), so that the corresponding tensor algebra \(T^1A\) can be defined as in Section 25.3. Let \(I_A\) be the two-sided ideal in \(T^1A\) generated by elements of the form

\[
[x, y]_A = x \otimes y - y \otimes x,
\]

where \(x, y \in A\). Of course, this uses the obvious embeddings of \(A = T^1A\) and \(T^2A\) in \(T^1A\). Thus \(I_A\) consists of elements of \(T^1A\) that can be obtained by multiplying elements of the form (25.5.1) by other elements of \(T^1A\) on the left and right, and taking finite sums of such products.
CHAPTER 25. UNIVERSAL ENVELOPING ALGEBRAS

The corresponding quotient

\[ UA = (TA)/\mathcal{I}_A \] (25.5.2)

is an associative algebra over \( k \). Let \( q_A \) be the canonical quotient mapping from \( TA \) onto \( UA \). Remember that \( 1 \in k = T^0A \) is the multiplicative identity element in \( TA \), so that \( e_{UA} = q_A(1) \) is the multiplicative identity element in \( UA \).

Let \( i = i_{UA} \) be the composition of the natural inclusion mapping of \( A = T^1A \) into \( TA \) with \( q_A \). Let us check that \( (UA, i) \) is a universal enveloping algebra of \( A \), as in Theorem 1.2 on p11 of [24], and as mentioned on p91 of [13].

It is easy to see that \( i \) is a Lie algebra homomorphism from \( A \) into \( UA \), by construction. Let \( B \) be an associative algebra over \( k \) with a multiplicative identity element \( e_B \), and let \( \phi \) be a Lie algebra homomorphism from \( A \) into \( B \). In particular, \( \phi \) is a homomorphism from \( A \) into \( B \) as modules over \( k \), and so there is a unique algebra homomorphism \( \rho \) from \( TA \) into \( B \) that is equal to \( \phi \) on \( A = T^1A \) and sends \( 1 \in k = T^0A \) to \( e_B \), as in Section 25.3. If \( x, y \in A \), then one can check that \( 25.5.1 \) is in the kernel of \( \rho \), because \( \phi \) is a Lie algebra homomorphism from \( A \) into \( B \). This implies that \( \mathcal{I}_A \) is contained in the kernel of \( \rho \), so that \( \rho \) can be expressed as

\[ \rho = \psi \circ q_A \] (25.5.3)

on \( TA \) for some algebra homomorphism \( \psi \) from \( UA \) into \( B \). Note that \( \psi(e_{UA}) = \psi(q_A(1)) = \rho(1) = e_B \). We can compose both sides of \( 25.5.3 \) with the natural inclusion of \( A = T^1A \) into \( TA \) to get that

\[ \phi = \psi \circ i \] (25.5.4)

on \( A \).

Suppose that \( \tilde{\psi} \) is another algebra homomorphism from \( UA \) into \( B \) such that \( \psi(e_{UA}) = e_B \) and \( \phi = \tilde{\psi} \circ i \). This implies that \( \tilde{\rho} = \tilde{\psi} \circ q_A \) is an algebra homomorphism from \( TA \) into \( B \) such that \( \tilde{\rho}(1) = \tilde{\psi}(q_A(1)) = \psi(e_{UA}) = e_B \) and \( \tilde{\rho} = \phi \) on \( A = T^1A \). It follows that \( \tilde{\rho} = \rho \) on \( TA \), as in Section 25.3. This means that \( \tilde{\psi} = \psi \) on \( UA \), so that \( \psi \) is unique. This could also be obtained from the fact that \( UA \) is generated by \( e_{UA} \) and \( i(A) \), as an algebra over \( k \), because \( TA \) is generated by the images of \( k = T^0A \) and \( A = T^1A \) in \( TA \), as an algebra over \( k \).

25.6 Some properties of \( UA \)

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( A \) be a module over \( k \). Also let \( q_0 \) be the mapping from \( TA \) into \( k \) that sends each element of \( TA \) to its component in \( T^0A = k \). This is an algebra homomorphism from \( TA \) onto \( k \). The kernel of \( q_0 \) is the two-sided ideal \( \mathcal{I}_0 \) in \( TA \) consisting of elements of \( TA \) whose component in \( T^0A \) is equal to 0, which corresponds to the direct sum of \( T^nA \) over \( n \geq 1 \) in \( TA \).
25.7. COMMUTING LIE SUBALGEBRAS

Suppose now that $(A, [\cdot, \cdot]_A)$ is a Lie algebra over $k$. Let $\mathcal{I}_A$ be the two-sided ideal in $TA$ associated to $A$ as in the previous section, and observe that

$$\mathcal{I}_A \subseteq \mathcal{I}_0.$$  

(25.6.1)

Remember that $q_A$ is the canonical quotient mapping from $TA$ onto $UA = (TA)/\mathcal{I}_A$, as before. Using (25.6.1), we get that the restriction of $q_A$ to $T^0A = k$ is injective, as mentioned on p91 of [13]. We can also express $q_0$ as the composition of $q_A$ with an algebra homomorphism from $UA$ onto $k$. Alternatively, the map from $A$ into $k$ that sends every element of $A$ to 0 is a Lie algebra homomorphism. This leads to an algebra homomorphism from $UA$ into $k$ that sends $e$ to 1.

Let $V$ be a module over $k$, and remember that the space $\text{Hom}_k(V, V)$ of module homomorphisms from $V$ into itself is an associative algebra over $k$ with respect to composition of mappings, and with the identity mapping on $V$ as the multiplicative identity element. A Lie algebra homomorphism from $A$ into $\text{Hom}_k(V, V)$ is the same as a representation of $A$, as a Lie algebra over $k$, on $V$. This leads to an algebra homomorphism from $UA$ into $\text{Hom}_k(V, V)$ that sends $e_{UA}$ to the identity mapping on $V$, and whose composition with $i_{UA}$ is the given representation of $A$ on $V$. Thus, if $V$ is a module over $A$ as a Lie algebra, then $V$ may be considered as a left module over $UA$ as an associative algebra, as in the remark on p11f of [24].

There is an exercise on p12 of [24], which is attributed to Bergman, and which asks one to show that $UA = k$ if and only if $A = \{0\}$. More precisely, $UA = k$ means that $UA = q_A(k)$, which happens exactly when $\mathcal{I}_A = \mathcal{I}_0$. Of course, if $A = \{0\}$, then $\mathcal{I}_0 = \{0\}$, and $\mathcal{I}_A = \mathcal{I}_0$. Conversely, let us check that $A = \{0\}$ when $\mathcal{I}_A = \mathcal{I}_0$. The exercise comes with the hint that one should use the adjoint representation. Remember that $A$ may be considered as a module over itself, as a Lie algebra over $k$, using the adjoint representation. Using this, we may consider $A$ as a left module over $UA$, as an associative algebra over $k$, as in the preceding paragraph. If $\mathcal{I}_A = \mathcal{I}_0$, then it follows that the action of any element of $A$ on $A$ by the adjoint representation is equal to 0, so that $A$ is commutative as a Lie algebra. In this case, the condition that $\mathcal{I}_A = \mathcal{I}_0$ implies that $A = T^1A = \{0\}$, as desired.

Let $A_0$ be a Lie subalgebra of $A$, so that $A_0$ also has a universal enveloping algebra $UA_0$, with multiplicative identity element $e_0 = e_{UA_0}$ and corresponding Lie algebra homomorphism $i_0 = i_{UA_0}$ from $A_0$ into $UA_0$. Consider the restriction $\phi_0$ of $i_{UA}$ to $A_0$, which is a Lie algebra homomorphism from $A_0$ into $UA$. This leads to a unique algebra homomorphism $\psi_0$ from $UA_0$ into $UA$ such that $\psi_0(e_0) = e_{UA}$ and $\psi_0 \circ i_0 = \phi_0$.

25.7 Commuting Lie subalgebras

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. Also let $UA$ be a universal enveloping algebra of $A$, with multiplicative identity element $e = e_{UA}$ and corresponding Lie algebra
homomorphism \( i = i_{UA} \) from \( A \) into \( UA \). Suppose that \( A_1 \) and \( A_2 \) are Lie subalgebras of \( A \), and let \( UA_1, UA_2 \) be universal enveloping algebras of \( A_1, A_2 \), respectively, with multiplicative identity elements \( e_i = e_{UA_i} \) and corresponding Lie algebra homomorphisms \( i_j = i_{UA_j} \) from \( A_j \) into \( UA \) for \( j = 1, 2 \). As in the previous section, the restriction \( \phi_j \) of \( i \) to \( A_j \) is a Lie algebra homomorphism from \( A_j \) into \( UA \) for \( j = 1, 2 \), which leads to a unique algebra homomorphism \( \psi_j \) from \( UA_j \) into \( UA \) such that \( \psi_j(e_j) = e \) and \( \psi_j \circ i_j = \phi_j \).

Suppose that \( A_1 \) and \( A_2 \) commute in \( A \), in the sense that

\[
[a_1, a_2]|_A = 0
\]

for every \( a_1 \in A_1 \) and \( a_2 \in A_2 \). This implies that

\[
i(a_1) i(a_2) = i(a_2) i(a_1)
\]

in \( UA \) for every \( a_1 \in A_1 \) and \( a_2 \in A_2 \), because \( i \) is a Lie algebra homomorphism from \( A \) into \( UA \). Using this, we get that

\[
\psi_1(x_1) \psi_2(x_2) = \psi_2(x_2) \psi_1(x_1)
\]

in \( UA \) for every \( x_1 \in UA_1 \) and \( x_2 \in UA_2 \). More precisely, this follows from (25.7.2) when \( x_j \in i_j(A_j) \) for \( j = 1, 2 \). Of course, (25.7.3) holds automatically when \( x_1 = e_1 \) or \( x_2 = e_2 \). It follows that (25.7.3) holds for every \( x_1 \in UA_1 \) and \( x_2 \in UA_2 \), because \( UA_j \) is generated by \( e_j \) and \( i_j(A_j) \) as an algebra over \( k \) for \( j = 1, 2 \). Let \( \psi \) be the unique linear mapping from \((UA_1) \otimes (UA_2)\) into \( UA \) such that

\[
\psi(x_1 \otimes x_2) = \psi_1(x_1) \psi_2(x_2)
\]

for every \( x_1 \in UA_1 \) and \( x_2 \in UA_2 \). This is an algebra homomorphism, as in Section 25.2, with \( \psi(e_1 \otimes e_2) = \psi_1(e_1) \psi_2(e_2) = e \).

Suppose now that \( A \) corresponds to the direct sum of \( A_1 \) and \( A_2 \), so that every element of \( A \) can be expressed in a unique way as a sum of elements of \( A_1 \) and \( A_2 \). Observe that

\[
a_1 \mapsto i_1(a_1) \otimes e_2
\]

is a Lie algebra homomorphism from \( A_1 \) into \((UA_1) \otimes (UA_2)\), and similarly that

\[
a_2 \mapsto e_1 \otimes i_2(a_2)
\]

is a Lie algebra homomorphism from \( A_2 \) into \((UA_1) \otimes (UA_2)\). If \( a_1 \in A_1 \) and \( a_2 \in A_2 \), then put

\[
\rho(a_1 + a_2) = i_1(a_1) \otimes e_2 + e_1 \otimes i_2(a_2),
\]

which defines a mapping from \( A \) into \((UA_1) \otimes (UA_2)\). It is easy to see that \( \rho \) is a Lie algebra homomorphism from \( A \) into \((UA_1) \otimes (UA_2)\), using (25.7.1). This implies that there is a unique algebra homomorphism \( \theta \) from \( UA \) into \((UA_1) \otimes (UA_2)\) such that \( \theta(e) = e_1 \otimes e_2 \) and \( \rho = \theta \circ i \).
Let us check that $\psi \circ \theta$ is the identity mapping on $U A$. Of course, $(\psi \circ \theta)(e) = \psi(\theta(e)) = \psi(e_1 \otimes e_2) = e$. If $a_1 \in A_1$ and $a_2 \in A_2$, then

\[
(25.7.8) \quad (\psi \circ \theta)(i(a_1 + a_2)) = \psi(\theta(a_1 + a_2)) \\
= \psi(i_1(a_1) \otimes e_2 + \psi(e_1 \otimes i_2(a_2))) \\
= \psi(\psi_1(a_1)) \psi_2(e_2) + \psi_1(e_1) \psi_2(i_2(a_2)) \\
= \phi_1(a_1) + \phi_2(a_2) = i(a_1 + a_2).
\]

This implies that $\psi \circ \theta$ is the identity mapping on $U A$, because $U A$ is generated as an algebra over $k$ by $e$ and $i(A)$.

Similarly, we would like to verify that $\theta \circ \psi$ is the identity mapping on $(U A_1) \otimes (U A_2)$. It suffices to show that

\[
(25.7.9) \quad (\theta \circ \psi)(x_1 \otimes x_2) = x_1 \otimes x_2
\]

for every $x_1 \in U A_1$ and $x_2 \in U A_2$. Observe that

\[
(25.7.10) \quad (\theta \circ \psi)(x_1 \otimes x_2) = \theta(\psi(x_1 \otimes x_2)) \\
= \theta(\psi_1(x_1) \psi_2(x_2)) = \theta(\psi_1(x_1)) \theta(\psi_2(x_2))
\]

for every $x_1 \in U A_1$ and $x_2 \in U A_2$.

Let us check that

\[
(25.7.11) \quad \theta(\psi_1(x_1)) = x_1 \otimes e_2
\]

for every $x_1 \in U A_1$. Of course, $\theta(\psi_1(e_1)) = \theta(e) = e_1 \otimes e_2$, so that (25.7.11) holds when $x_1 = e_1$. If $a_1 \in A_1$, then

\[
(25.7.12) \quad \theta(\psi_1(i_1(a_1))) = \theta(i(a_1)) = \rho(a_1) = i_1(a_1) \otimes e_2.
\]

This shows that (25.7.11) holds when $x_1 \in i_1(A_1)$. It follows that (25.7.11) holds for every $x_1 \in U A_1$, because $U A_1$ is generated by $e_1$ and $i_1(A_1)$, as an algebra over $k$.

If $x_2 \in U A_2$, then

\[
(25.7.13) \quad \theta(\psi_2(x_2)) = e_1 \otimes x_2,
\]

for the same reasons. If $x_1 \in U A_1$ as well, then we get that

\[
(25.7.14) \quad (\theta \circ \psi)(x_1 \otimes x_2) = \theta(\psi_1(x_1)) \theta(\psi_2(x_2)) \\
= (x_1 \otimes e_2)(e_1 \otimes x_2) = x_1 \otimes x_2,
\]

as desired. Thus $U A$ is isomorphic to $(U A_1) \otimes (U A_2)$ in a natural way, as on p12 of [24].

## 25.8 Symmetric algebras

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be a module over $k$. We may consider $A$ as a commutative Lie algebra over $k$,
by putting \([x,y]_A = 0\) for every \(x,y \in A\). In this case, a universal enveloping algebra of \(A\) as a Lie algebra over \(k\) is called a symmetric algebra of \(A\) as a module over \(k\), and is denoted \(SA\), as on p12 of [24].

More precisely, a symmetric algebra \(SA\) of \(A\) comes with a multiplicative identity element \(e = e_{SA}\), and a Lie algebra homomorphism \(i = i_{SA}\) from \(A\) into \(SA\). In this situation, the condition that \(i\) be a Lie algebra homomorphism means that \(i\) is a homomorphism from \(A\) into \(SA\), as modules over \(k\), and that

\[
i(x)i(y) = i(y)i(x)
\]

for every \(x,y \in A\). This implies that \(SA\) is commutative as an algebra over \(k\), because it is generated by \(i(A)\) and \(e\), as before.

The two-sided ideal \(I_A\) in the tensor algebra \(TA\) defined in Section 25.5 is generated by elements of the form

\[
x \otimes y - y \otimes x
\]

with \(x,y \in A\) in this situation. This means that \(I_A\) consists of elements of \(TA\) that can be obtained by multiplying elements of the form (25.8.2) by other elements of \(TA\) on the left and right, and taking finite sums of such elements, as before. Thus a symmetric algebra of \(A\) can be obtained as the quotient

\[
SA = (TA)/I_A,
\]

as in Section 25.5. This corresponds to the definition of the symmetric algebra on p89 of [13], and the characterization as a universal enveloping algebra of a commutative Lie algebra is mentioned in the example on p91 of [13].

Let us identify \(T^nA\) with a submodule of \(TA\), as a module over \(k\), in the obvious way. It is easy to see that \(I_A\) corresponds to the direct sum of \(I_A \cap (T^nA)\) over \(n \geq 2\) in \(TV\). More precisely, \(I_A \cap (T^nA)\) consists of finite sums of elements that can be expressed as the product of the form (25.8.2) by an element of \(T^jA\) on the left and \(T^lA\) on the right, where \(j, l\) as nonnegative integers such that \(j + l + 2 = n\). In particular, \(I_A \cap (T^nA) = \{0\}\) when \(n = 0\) or \(1\). If \(n\) is any nonnegative integer, then \(I_A \cap (T^nA)\) is a submodule of \(T^n\), as a module over \(k\), so that the quotient

\[
S^nA = (T^nA)/(I_A \cap (T^nA))
\]

can be defined as a module over \(k\). Thus (25.8.3) may be considered as the direct sum of \(S^nA\) over \(n \geq 0\), as a module over \(k\). Note that \(S^0A = T^0A = k\), and \(S^1A = T^1A = A\).

Let \(SA\) be a symmetric algebra of \(A\) again, with \(e = e_{SA}\) and \(i = i_{SA}\) as before. The second property in the definition of a universal enveloping algebra says that if \(B\) is an associative algebra over \(k\) with a multiplicative identity element \(e_B\), and \(\phi\) is a Lie algebra homomorphism from \(A\) into \(B\), then there is a unique algebra homomorphism \(\psi\) from \(SA\) into \(B\) such that \(\psi(e) = e_B\) and \(\psi \circ i = \phi\). In this situation, a mapping \(\phi\) from \(A\) into \(B\) is a Lie algebra homomorphism if \(\phi\) is a homomorphism from \(A\) into \(B\), as modules over \(k\), and

\[
\phi(x)\phi(y) = \phi(y)\phi(x)
\]
25.9. A FILTRATION ON $UA$

for every $x, y \in A$. Of course, (25.8.5) holds automatically when $B$ is commutative, in which case this corresponds to a remark on p90 of [13]. Remember that $SA$ is unique up to isomorphism, as in Section 25.4.

Suppose now that $A$ is free as a module over $k$, so that $A$ corresponds to the direct sum of a nonempty family of copies of $k$. Equivalently, there is a nonempty set $I$ and an element $e_j$ of $A$ for each $j \in I$ so that every element of $A$ can be expressed in a unique way as a finite linear combinations of $e_j$’s, with coefficients in $k$. Let $X_j, j \in I$, be a commuting family of indeterminates, so that the corresponding polynomial algebra $k[[X_j : j \in I]]$ can be defined as in Section 25.1. There is a unique homomorphism $i$ from $A$ into $k[[X_j : j \in I]]$, as modules over $k$, such that

$$i(e_j) = X_j$$

for every $j \in I$. Let us check that this satisfies the requirements of a symmetric algebra of $A$.

Note that $i$ is a Lie algebra homomorphism from $A$ into $k[[X_j : j \in I]]$, because $k[[X_j : j \in I]]$ is a commutative algebra over $k$. Let $B$ be an associative algebra over $k$ with a multiplicative identity element $e_B$, and let $\phi$ be a Lie algebra homomorphism from $A$ into $B$, where $A$ is considered as a commutative Lie algebra over $k$. This means that $\phi$ is a homomorphism from $A$ into $B$, as modules over $k$, that satisfies (25.8.5), as before. Put

$$b_j = \phi(e_j)$$

for each $j \in I$, so that $b_jb_l = b_lb_j$ for every $j, l \in I$. Let $b$ be the element of the Cartesian power $B^I$ whose $j$th coordinate is equal to $b_j$ for each $j \in I$. If $f(X) \in k[[X_j : j \in I]]$, then put

$$\psi(f(X)) = f(b),$$

where the right side is as in Section 25.1. Thus $\psi$ is an algebra homomorphism from $k[[X_j : j \in I]]$ into $B$, as before, and it is easy to see that $\psi \circ i = \phi$ on $A$. Of course, $\psi(1) = e_B$, by construction, and $\psi$ is uniquely determined by these properties. This corresponds to some remarks on p12 of [24], and is mentioned on p90 of [13] as well.

25.9 A filtration on $UA$

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. Also let $UA$ be a universal enveloping algebra of $A$, with multiplicative identity element $e_{UA}$ and Lie algebra homomorphism $i_{UA}$ from $A$ into $UA$. If $n$ is a nonnegative integer, then we let $U_nA$ be the submodule of $UA$, as a module over $k$, generated by products of the form

$$i_{UA}(x_1) \cdots i_{UA}(x_m),$$

where $x_1, \ldots, x_m \in A$ and $m \leq n$. We interpret $e_{UA}$ as being such a product with $m = 0$, so that multiples of $e_{UA}$ by elements of $k$ are in $U_nA$ for every
CHAPTER 25. UNIVERSAL ENVELOPING ALGEBRAS

\( n \geq 0 \). More precisely, \( U_0A \) consists exactly of multiples of \( e_{UA} \) by elements of \( k \), and

\[ U_nA \subseteq U_{n+1}A \]

for every \( n \geq 0 \). Note that \( \bigcup_{n=0}^{\infty} U_nA = UA \). This is the way that \( U_nA \) is defined on p13 of [24].

Alternatively, we can take \( UA \) to be obtained from the tensor algebra \( TA \) as in Section 25.5. If \( n \) is a nonnegative integer, then let \( T_nA \) be the submodule of \( TA \) that corresponds to the direct sum of \( T^mA \) over \( m = 0, \ldots, n \). Thus \( T_0A \) corresponds to \( T^0A \), as a submodule of \( TA \), \( T_nA \subseteq T_{n+1}A \) for every \( n \geq 0 \), and \( TA = \bigcup_{n=0}^{\infty} T_nA \). If \( UA \) is the quotient of \( TA \) discussed before, then \( U_nA \) corresponds to the image of \( T_nA \) under the quotient mapping from \( TA \) onto \( UA \). This is how \( U_nA \) is defined on p91 of [13].

If \( n \) is a positive integer, then put

\[ gr_nUA = (U_nA)/(U_{n-1}A), \]

where the quotient is defined as a module over \( k \). Also put \( gr_0UA = U_0A \), and

\[ gr UA = \bigoplus_{n=0}^{\infty} gr_nUA, \]

considered as a module over \( k \) for the moment. If \( n, r \) are nonnegative integers, then multiplication on \( UA \) defines a bilinear mapping from \((U_nA) \times (U_rA)\) into \( U_{n+r}A \). It is easy to see that this determines a bilinear mapping from \((gr_nUA) \times (gr_rUA)\) into \( gr_{n+r}UA \) in a natural way. This defines a bilinear mapping of multiplication from \((gr UA) \times (gr UA)\) into \( gr UA \). One can check that multiplication on \( gr UA \) is associative, and that the element of \( gr UA \) corresponding to \( e_{UA} \in U_0A = gr_0UA \) is the multiplicative identity element. This is called the \textit{graded algebra} associated to \( UA \), as on p13 of [24].

Note that \( i_{UA} \) maps \( A \) into \( U_1A \), which is mapped into \( gr_1UA \) by the quotient mapping. This gives a natural homomorphism from \( A \) into \( gr UA \), as modules over \( k \). One can check that \( gr UA \) is generated, as an algebra over \( k \), by the image of \( A \) under this mapping and the multiplicative identity element, as in Proposition 4.1 on p13 of [24].

Theorem 4.2 on p13 of [24] states that \( gr UA \) is commutative as an algebra over \( k \). To see this, let \( x, y \in A \) be given, and remember that

\[ i_{UA}(x) i_{UA}(y) - i_{UA}(y) i_{UA}(x) = i_{UA}([x, y]_A) \]

in \( UA \), because \( i_{UA} \) is a Lie algebra homomorphism from \( A \) into \( UA \). Of course, \( i_{UA}(x) i_{UA}(y) \) and \( i_{UA}(y) i_{UA}(x) \) are elements of \( U_2A \), which are mapped to the same element of \( gr_2UA \), by (25.9.5). This means that the images of \( x, y \) in \( gr UA \) commute. It follows that \( gr UA \) is commutative, because \( gr UA \) is generated as an algebra by the image of \( A \) and the multiplicative identity element, as in the preceding paragraph.

Let \( SA \) be a symmetric algebra of \( A \), as a module over \( k \), with multiplicative identity element \( e_{SA} \), and homomorphism \( i_{SA} \) from \( A \) into \( SA \), as modules over
k. Because $\text{gr} U A$ is commutative, there is a unique algebra homomorphism $\rho$ from $S A$ into $\text{gr} U A$ such that $\rho(e_{S A})$ is the multiplicative identity element in $\text{gr} U A$, and $\rho \circ i_{S A}$ is the natural mapping from $A$ into $\text{gr} U A$. More precisely, $\rho$ maps $S A$ onto $\text{gr} U A$, because $\text{gr} U A$ is generated as an algebra by the image of $A$ and the multiplicative identity element, as mentioned on p13 of [24].

This corresponds to the lemma on p91 of [13], which is formulated with $U A$ given by the quotient of $T A$ as in Section 25.5. If $n$ is a nonnegative integer, then there is a natural module homomorphism from $T^n A$ into $U^n A$, using the quotient mapping from $T A$ onto $U A$. More precisely, this mapping can be obtained by first mapping $T^n A$ into $T^n U A$, and mapping $T^n U A$ onto $U^n A$ using the quotient mapping from $T A$ onto $U A$. This leads to a module homomorphism from $T^n A$ into $\text{gr}_n U A$, by composition with the quotient mapping from $U^n A$ onto $\text{gr}_n U A$. In fact, this module homomorphism maps $T^n A$ onto $\text{gr}_n U A$. We can combine these module homomorphisms to get a module homomorphism from $T A$ onto $\text{gr} U A$. This maps the multiplicative identity element in $T A$ to the multiplicative identity element in $\text{gr} U A$. One can check that this is an algebra homomorphism from $T A$ onto $\text{gr} U A$.

Remember that a symmetric algebra $S A$ of $A$ can be obtained as a quotient of $T A$ as well. One can verify that the ideal in $T A$ used to get $S A$ is contained in the kernel of the homomorphism from $T A$ onto $\text{gr} U A$ mentioned in the preceding paragraph, using (25.9.5). This leads to an algebra homomorphism from $S A$ onto $\text{gr} U A$, as before.

25.10 The Poincaré–Birkhoff–Witt theorem

Let us continue with the same notation and hypotheses as in the previous section, and suppose also that $A$ is a free module over $k$. The Poincaré–Birkhoff–Witt theorem states that the homomorphism $\rho$ from $S A$ onto $\text{gr} U A$ defined in the previous section is an isomorphism, as in Theorem 4.3 on p14 of [24]. Of course, if $k$ is a field, so that $A$ is a vector space over $k$, then $A$ is automatically free as a module over $k$. This corresponds to the formulation of the theorem on p92 of [13].

To say that $A$ is a free module over $k$ means that there is a nonempty set $I$ and a family \( \{x_j\}_{j \in I} \) of elements of $A$ indexed by $I$ such that every element of $A$ can be expressed in a unique way as a finite linear combination of the $x_j$’s, with coefficients in $k$. We may refer to \( \{x_j\}_{j \in I} \) as a basis for $A$, as a free module over $k$. Let $\preceq$ be a linear ordering on $I$.

Let $n$ be a nonnegative integer, and let $j_1, \ldots, j_m$ be $m$ elements of $I$ for some $m \leq n$, with

\[
(25.10.1) \quad j_1 \preceq \cdots \preceq j_m.
\]

Thus the product

\[
(25.10.2) \quad i_{U A}(x_{j_1}) \cdots i_{U A}(x_{j_m})
\]

is an element of $U^n A$. As before, we interpret (25.10.2) as being equal to $e_{U A}$ when $m = 0$, in which case (25.10.1) is considered to hold vacuously. Lemma
CHAPTER 25. UNIVERSAL ENVELOPING ALGEBRAS

4.4 on p14 of [24] states that $U_nA$ is generated, as a module over $k$, by products of the form (25.10.2).

The proof is by induction on $n$, with the base case $n = 0$ being trivial. Suppose that $n \geq 1$, and remember that $U_nA$ is generated as a module over $k$ by products of the form (25.10.2) without the condition (25.10.1), by definition. Using (25.9.5), we can express an element of $U_nA$ as a linear combination of terms of the form (25.10.2), with coefficients in $k$, $m = n$, and the condition (25.10.1), plus an element of $U_{n-1}A$. This implies that $U_nA$ is generated as a module over $k$ in the desired way, because of the corresponding property for $U_{n-1}A$, by the induction hypothesis.

Consider the following statement:

(25.10.3) the family of products of the form (25.10.2) with (25.10.1) and $m \geq 0$ is a basis for $UA$ as a module over $k$.

Lemma 4.5 on p14 of [24] asserts that this is equivalent to the Poincaré–Birkhoff–Witt theorem. Lemma C on p92 of [13] corresponds to the fact that the theorem implies (25.10.3).

If $m \in \mathbb{Z}_+$, then let $\mathcal{M}_m(I)$ be the set of $m$-tuples $M = (j_1, \ldots, j_m)$ of elements of $I$ that satisfy (25.10.1). We consider $M = \emptyset$ to be the unique element of $\mathcal{M}_0(I)$, and put

(25.10.4) $\mathcal{M}(I) = \bigcup_{m=0}^{\infty} \mathcal{M}_m(I)$.

If $M \in \mathcal{M}_m(I)$ for some $m \geq 0$, then $m$ is called the length of $M$, and may be denoted $\text{length}(M)$.

Remember that $\mathcal{M}_I$ is the set of all multi-indices with respect to $I$, as in Section 25.1. If $\alpha \in \mathcal{M}_I$, then let $\mathcal{M}_\alpha \in \mathcal{M}(I)$ be the $m$-tuple with $m = |\alpha|$ obtained by taking the $j \in I$ with $\alpha(j) \neq 0$ in order and with multiplicity $\alpha(j)$, so that $M_\alpha = \emptyset$. Similarly, if $M \in \mathcal{M}_m(I)$ for some $m \geq 0$, then we can define $\alpha_M \in \mathcal{M}_I$ by taking $\alpha_M(j)$ to be the number of coordinates of $M$ equal to $j$ for each $j \in I$. This defines a one-to-one correspondence between $\mathcal{M}(I)$ and $\mathcal{M}_I$.

Let $X_j$, $j \in I$, be a commuting family of indeterminates, with the corresponding polynomial algebra $k[\{X_j : j \in I\}]$, as in Section 25.1. There is a unique homomorphism from $A$ into $k[\{X_j : j \in I\}]$, as modules over $k$, that sends $x_j \in A$ to $X_j$ for every $j \in I$. Using this module homomorphism, $k[\{X_j : j \in I\}]$ satisfies the requirements of a symmetric algebra of $A$, as in Section 25.8. Of course, the collection of monomials $X^n$, $\alpha \in \mathcal{M}_I$, is a basis for $k[\{X_j : j \in I\}]$ as a module over $k$.

If $M = (j_1, \ldots, j_n) \in \mathcal{M}_n(I)$ for some $n \geq 0$, then put

(25.10.5) $x_M = i_{UA}(x_{j_1}) \cdots i_{UA}(x_{j_n})$,

which is interpreted as being $e_{UA}$ when $M = \emptyset$, as before. This is an element of $U_nA$, and we let $\overline{x_M}$ be its image in $\text{gr}_nUA$ under the corresponding quotient.
mapping. Using \( k[[X_j : j \in I]] \) as the symmetric algebra of \( A \), the mapping \( \rho \) from \( SA \) onto \( gr\ UA \) defined in the previous section corresponds exactly to the module homomorphism from \( k[[X_j : j \in I]] \) into \( gr\ UA \) with

\[
X^\alpha \mapsto \sum_{M \in M_n(I)} c_M x_M
\]

for every \( \alpha \in \mathcal{M}_I \).

The Poincaré–Birkhoff–Witt theorem is equivalent to the statement that \( \rho \) is injective, because we already know that it is a surjective algebra homomorphism. The injectivity of \( \rho \) is equivalent to saying that if \( n \geq 0 \), \( c_M \in k \) for every \( M \in \mathcal{M}_n(I) \), and

\[
\sum_{M \in \mathcal{M}_n(I)} c_M x_M = 0
\]

in \( gr\ UA \), then \( c_M = 0 \) for every \( M \in \mathcal{M}_n(I) \). This holds trivially when \( n = 0 \), and so we may as well take \( n \geq 1 \). Note that (25.10.7) is the same as saying that

\[
\sum_{M \in \mathcal{M}_n(I)} c_M x_M \equiv 0 \mod U_{n-1} A
\]

in \( U_n A \) when \( n \geq 1 \). This means that there are \( c_M \in k \) for \( M \in \mathcal{M}(I) \) with \( \text{length}(M) \leq n - 1 \) such that

\[
\sum_{M \in \mathcal{M}_n(I)} c_M x_M = \sum_{M \in \mathcal{M}(I) \atop \text{length}(M) < n} c_M x_M
\]

in \( U_n A \), because \( U_{n-1} A \) is generated as a module over \( k \) by the \( x_M \) with \( M \) in \( \mathcal{M}(I) \) and \( \text{length}(M) \leq n - 1 \), as before.

Of course, (25.10.3) is the same as saying that the family of \( x_M, M \in \mathcal{M}(I) \), is a basis for \( UA \) as a module over \( k \). If this holds, then (25.10.9) implies that \( c_M = 0 \) for every \( M \in \mathcal{M}(I) \) with \( \text{length}(M) \leq n \), and in particular when \( \text{length}(M) = n \), as desired. Conversely, if \( \rho \) is injective, then (25.10.9) implies that \( c_M = 0 \) for every \( M \in \mathcal{M}(I) \) with \( \text{length}(M) = n \), as in the preceding paragraph. In this case, one can repeat the process to get that \( c_M = 0 \) for every \( M \in \mathcal{M}(I) \) with \( \text{length}(M) \leq n \). This implies that the family of \( x_M \), \( M \in \mathcal{M}(I) \) is a basis for \( UA \) as a module over \( k \), because \( UA \) is generated by the \( x_M \)’s, \( M \in \mathcal{M}(I) \), as a module over \( k \).

25.11 The rest of the proof

Let us continue with the same notation and hypotheses as in the previous two sections. In order to prove the Poincaré–Birkhoff–Witt theorem, we take \( \preceq \) to be a well-ordering on \( I \). Let \( V \) be a free module over \( k \) with a basis consisting of elements \( Z_M \) for each \( M \in \mathcal{M}(I) \), as on p14 of [24]. As a module over \( k \), \( V \) is isomorphic to \( k[\{X_j : j \in I\}] \), which is a symmetric algebra of \( A \), and the argument on p93 of [13] uses this.
If \( j \in I \) and \( M = (j_1, \ldots, j_n) \in \mathcal{M}(I) \), then we put \( j \preceq M \) when \( j \preceq j_1 \). In this case, we put
\[
(25.11.1) \quad jM = (j, j_1, \ldots, j_n),
\]
which is an element of \( \mathcal{M}(I) \) too. If \( M = \emptyset \), then \( j \preceq M \) is interpreted as holding automatically, and \( jM \) has only the one coordinate \( j \). The main lemma on p14 of [24] states that \( V \) can be made a module over \( A \), as a Lie algebra over \( k \), in such a way that
\[
(25.11.2) \quad x_j \cdot Z_M = Z_{jM}
\]
for every \( j \in I \) and \( M \in \mathcal{M}(I) \) with \( j \preceq M \). This corresponds to Lemma B on p94 of [13], part (a) in particular.

We first want to define \((x, v) \mapsto x \cdot v \) as a bilinear mapping from \( A \times V \) into \( V \). To do this, it suffices to define \( x_j \cdot Z_M \) as an element of \( V \) for every \( j \in I \) and \( M \in \mathcal{M}(I) \). More precisely, we want to do this in such a way that
\[
(25.11.3) \quad x_j \cdot Z_M \text{ is a finite linear combination of } Z_{l'} \text{'s with coefficients in } k, \ L \in \mathcal{M}(I), \text{ and } \text{length}(L) \leq \text{length}(M) + 1.
\]

Note that (25.11.2) has this property.

Let \( j \in I \) and \( M \in \mathcal{M}(I) \) be given. Suppose that \( x_{j_0} \cdot Z_N \in V \) has already been defined and satisfies (25.11.3) for \( j_0 \in I \) and \( N \in \mathcal{M}(I) \) such that either
\[
(25.11.4) \quad \text{length}(N) < \text{length}(M),
\]
or
\[
(25.11.5) \quad \text{length}(N) = \text{length}(M), \ j_0 \preceq j, \text{ and } j_0 \neq j.
\]
If \( j \preceq M \), then \( x_j \cdot Z_M \) can be defined as in (25.11.2). Otherwise, there are \( j_1 \in I \) and \( N \in \mathcal{M}(I) \) such that
\[
(25.11.6) \quad j_1 \preceq N, \ M = j_1 N, j_1 \preceq j, \text{ and } j_1 \neq j.
\]
In this case, we would like to define \( x_j \cdot Z_M \in V \) by
\[
(25.11.7) \quad x_j \cdot Z_M = x_{j_1} \cdot (x_j \cdot Z_N) + ([x_j, x_{j_1}]_A) \cdot Z_N.
\]
It is easy to see that the right side is defined as an element of \( V \) under these conditions, and that it satisfies (25.11.3) as well. One can use this to define \( x_j \cdot Z_M \) for every \( j \in I \) and \( M \in \mathcal{M}(I) \), because \( I \) is well ordered by \( \preceq \).

In order to show that \( V \) is a module over \( A \) as a Lie algebra over \( k \), we need to check that
\[
(25.11.8) \quad x \cdot (y \cdot v) - y \cdot (x \cdot v) = ([x, y]_A) \cdot v
\]
for every \( x, y \in A \) and \( v \in V \). To do this, it suffices to verify that
\[
(25.11.9) \quad x_j \cdot (x_l \cdot Z_N) - x_l \cdot (x_j \cdot Z_N) = ([x_j, x_l]_A) \cdot Z_N
\]
for every \( j, l \in I \) and \( N \in \mathcal{M}(I) \). Arguing by induction on \( \text{length}(N) \) and \( \min(j, l) \) as on p15 of [24], we may suppose that
\[
(25.11.10) \quad x_{j_0} \cdot (x_{l_0} \cdot Z_{N_0}) - x_{l_0} \cdot (x_{j_0} \cdot Z_{N_0}) = ([x_{j_0}, x_{l_0}]_A) \cdot Z_{N_0}
\]
holds for every \( j_0, l_0 \in I \) and \( N_0 \in \mathcal{M}(I) \) such that either
\[
\text{(25.11.11)} \quad \text{length} N_0 < \text{length} N
\]
or
\[
\text{(25.11.12)} \quad \text{length}(N_0) = \text{length}(N), \min(j_0, l_0) \leq \min(j, l), \quad \text{and} \quad \min(j_0, l_0) \neq \min(j, l).
\]

We may as well suppose that
\[
\text{(25.11.13)} \quad l \leq j \quad \text{and} \quad l \neq j,
\]
because both sides of (25.11.9) are antisymmetric in \( j \) and \( l \), and are equal to 0 when \( j = l \). Suppose first that \( l \leq N \), so that \( x_l \cdot Z_N = Z_l N \). Thus (25.11.7) holds with \( M = l N \) and \( j_1 = l \). This implies (25.11.9), as desired.

Otherwise, there are \( r \in I \) and \( L \in \mathcal{M}(I) \) such that
\[
\text{(25.11.14)} \quad r \leq L, \quad N = r L, \quad r \leq l, \quad \text{and} \quad r \neq l.
\]
This means that \( Z_N = Z_{rL} = x_r \cdot Z_l \), as in (25.11.2). In this case, (25.11.9) is the same as saying that
\[
\text{(25.11.15)} \quad x_j \cdot (x_l \cdot (x_r \cdot Z_L)) - x_l \cdot (x_j \cdot (x_r \cdot Z_L)) = ([x_j, x_l]_A) \cdot (x_r \cdot Z_L).
\]

Consider the analogous equations
\[
\text{(25.11.16)} \quad x_l \cdot (x_r \cdot (x_j \cdot Z_L)) - x_r \cdot (x_l \cdot (x_j \cdot Z_L)) = ([x_l, x_r]_A) \cdot (x_j \cdot Z_L)
\]
and
\[
\text{(25.11.17)} \quad x_r \cdot (x_j \cdot (x_l \cdot Z_L)) - x_j \cdot (x_r \cdot (x_l \cdot Z_L)) = ([x_r, x_j]_A) \cdot (x_l \cdot Z_L),
\]
where the indices \( j, l, \) and \( r \) have been cyclically permuted. These two equations hold by induction, because \( \min(l, r) = \min(r, j) = r \) is strictly less than \( \min(j, l) = l \). This also uses the fact that \( x_j \cdot Z_L \) and \( x_l \cdot Z_L \) can be expressed as finite linear combinations of \( Z_{N_0} \)'s with \( N_0 \in \mathcal{M}(I) \) and
\[
\text{(25.11.18)} \quad \text{length}(N_0) \leq \text{length}(L) + 1 = \text{length}(N),
\]
as in (25.11.3).

If \( x, y \in A \), then
\[
\text{(25.11.19)} \quad x \cdot (y \cdot Z_L) - y \cdot (x \cdot Z_L) = ([x, y]_A) \cdot Z_L,
\]
by induction, because
\[
\text{(25.11.20)} \quad \text{length} L = \text{length}(N) - 1 < \text{length}(N).
\]
This implies that

\[
\begin{align*}
([x_j, x_l]_A \cdot x_r \cdot Z_L) &= x_r \cdot ([x_j, x_l]_A \cdot Z_L) + ([x_j, x_l]_A \cdot x_r) \cdot Z_L \\
(25.11.21)
&= x_r \cdot (x_j \cdot (x_l \cdot Z_L)) - x_r \cdot (x_l \cdot (x_j \cdot Z_L)) \\
&+ ([x_j, x_l]_A, x_r) \cdot Z_L,
\end{align*}
\]

using (25.11.19) in both steps. The left side of (25.11.21) is the same as the right side of (25.11.15), so that (25.11.15) is equivalent to

\[
\begin{align*}
&x_j \cdot (x_l \cdot (x_r \cdot Z_L)) - x_l \cdot (x_j \cdot (x_r \cdot Z_L)) \\
&= x_r \cdot (x_j \cdot (x_l \cdot Z_L)) - x_r \cdot (x_l \cdot (x_j \cdot Z_L)) \\
&+ ([x_j, x_l]_A, x_r) \cdot Z_L.
\end{align*}
\]

Similarly, (25.11.16) and (25.11.17) imply that the equations analogous to (25.11.22) but with the indices \( j, l, \) and \( r \) permuted cyclically hold. Consider the equation obtained by adding the left and right sides of the three equations corresponding to (25.11.22) and the two analogous equations for (25.11.16) and (25.11.17). The terms on the left side of the new equation all cancel with terms on the right side of the new equation. There are also three terms on the right side of the new equation that correspond to the last term on the right side of (25.11.22), and the two analogous terms with the indices permuted cyclically. The sum of these three terms is equal to 0, by the Jacobi identity.

This means that the new equation holds automatically. Because the two equations analogous to (25.11.22) with the indices permuted cyclically hold, we get that (25.11.22) holds as well. It follows that (25.11.15) holds, so that (25.11.9) holds, as desired. This shows that \( V \) is a module over \( A \) as a Lie algebra over \( k \).

We may consider \( V \) as a left module over \( UA \) too, as an associative algebra over \( k \), as in Section 25.6. If \( M \in \mathcal{M}(I) \), then \( x_M \in UA \) can be defined as in (25.10.5). One can check that

\[
(25.11.23)

x_M \cdot Z_\emptyset = Z_M,
\]

using (25.11.2) and induction on the length of \( M \). In particular, if \( M = \emptyset \), then \( x_M \) is interpreted as being \( e_{UA} \), which acts on \( V \) by the identity mapping.

Suppose that \( c_M \) is an element of \( k \) for every \( M \in \mathcal{M}(I) \), with \( c_M = 0 \) for all but finitely many \( M \), and

\[
(25.11.24)

\sum_{M \in \mathcal{M}(I)} c_M x_M = 0
\]

in \( UA \). This implies that

\[
(25.11.25)

\sum_{M \in \mathcal{M}(I)} c_M Z_M = \sum_{M \in \mathcal{M}(I)} c_M x_M \cdot Z_\emptyset = 0
\]

in \( V \). It follows that \( c_M = 0 \) for every \( M \in \mathcal{M}(I) \), because the \( Z_M \)'s form a basis for \( V \) as a free module over \( k \), by construction. This means that the
family of $x_M$, $M \in \mathcal{M}(I)$, is a basis for $UA$ as a module over $k$, because $UA$ is generated by the $x_M$'s as a module over $k$, as in the previous section. This completes the proof of the Poincaré–Birkhoff–Witt theorem, as before.

## 25.12 Some corollaries

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot])$ be a Lie algebra over $k$. Also let $UA$ be a universal enveloping algebra of $A$, with multiplicative identity element $e = e_{UA}$ and Lie algebra homomorphism $i = i_{UA}$ from $A$ into $UA$. Remember that the submodule $U_nA$ of $UA$, as a module over $k$, is defined for nonnegative integers $n$ as in Section 25.9. In particular, $U_0A$ consists of multiples of $e$ by elements of $k$, and $U_1A$ consists of sums of elements of $U_0A$ and $i(A)$.

More precisely, $U_1A$ corresponds to the direct sum of $U_0A$ and $i(A)$, as a module over $k$, as indicated on p13 of [24]. This can be seen using the natural algebra homomorphism from $UA$ onto $k$ mentioned in Section 25.6. This implies that the restriction of the quotient mapping from $U_1A$ onto $gr_1UA = (U_1A)/(U_0A)$ to $i(A)$ is a one-to-one mapping from $i(A)$ onto $gr_1UA$.

If $A$ is free as a module over $k$, then the Poincaré–Birkhoff–Witt theorem implies that $i$ is an injective mapping from $A$ into $UA$. This is Corollary 1 on p16 of [24], which corresponds to Corollary B on p92 of [13].

Let $A_0$ be a Lie subalgebra of $A$, and let $UA_0$ be a universal enveloping algebra of $A_0$, with multiplicative identity element $e_0 = e_{UA_0}$ and Lie algebra homomorphism $i_0 = i_{UA_0}$ from $A_0$ into $UA_0$. As in Section 25.6, the restriction $\phi_0$ of $i$ to $A_0$ is a Lie algebra homomorphism from $A_0$ into $UA$, which leads to a unique algebra homomorphism $\phi_0$ from $UA_0$ into $UA$ such that $\psi_0(e_0) = e$ and $\psi_0 \circ i_0 = \phi_0$. Suppose that $A_0$ is free as a module over $k$, with basis $\{x_j\}_{j \in I_0}$ for some nonempty set $I_0$. Suppose also that $A$ is free as a module over $k$, and that $\{x_j\}_{j \in I_0}$ can be extended to a basis $\{x_j\}_{j \in I}$ of $A$, with $I_0 \subseteq I$. Let $\leq$ be a linear ordering on $I$, with $j_0 \leq l$ for every $j_0 \in I_0$ and $l \in I \setminus I_0$. Thus $\mathcal{M}(I)$ and $x_M \in UA$, $M \in \mathcal{M}(I)$, can be defined as in Section 25.10, and the family of $x_M$, $M \in I$, is a basis for $UA$, as a module over $k$, by the Poincaré–Birkhoff–Witt theorem. This implies that $\psi_0$ is an injective mapping from $UA_0$ into $UA$, and more precisely that $UA$ may be considered as a free module over $UA_0$, using $x_M$ with $M \in \mathcal{M}(I \setminus I_0)$. This corresponds to Corollary D on p92 of [13].

Let $A_1$ and $A_2$ be Lie subalgebras of $A$, and let $UA_1$, $UA_2$ be universal enveloping algebras of $A_1$, $A_2$, respectively, with multiplicative identity elements $e_j = e_{UA_j}$ and Lie algebra homomorphisms $i_j = i_{UA_j}$ from $A_j$ into $UA_j$ for $j = 1, 2$. As before, the restriction $\phi_j$ of $i$ to $A_j$ is a Lie algebra homomorphism from $A_j$ into $UA$ for $j = 1, 2$, which leads to a unique algebra homomorphism $\psi_j$ from $UA_j$ into $UA$ such that $\psi_j(e_j) = e$ and $\psi_j \circ i_j = \phi_j$. Observe that

\[
(25.12.1) \quad (w_1, w_2) \mapsto \psi_1(w_1) \psi_2(w_2)
\]

is a bilinear mapping from $(UA_1) \times (UA_2)$ into $UA$. Using this, we get a unique
linear mapping from \((UA_1) \otimes (UA_2)\) into \(UA\) with

\[
(25.12.2) \quad w_1 \otimes w_2 \mapsto \psi_1(w_1) \psi_2(w_2)
\]

for every \(w_1 \in UA_1\) and \(w_2 \in UA_2\).

Suppose that \(A\) corresponds to the direct sum of \(A_1\) and \(A_2\), as a module over \(k\), and that \(A_1\) and \(A_2\) are free as modules over \(k\). Let \(\{x_{ij}\}_{J \in I_1}\) and \(\{x_{ij}\}_{J \in I_2}\) be bases for \(A_1\) and \(A_2\), respectively, where \(I_1 \cap I_2 = \emptyset\). Of course, \(A\) is a free module over \(k\) as well, with basis \(\{x_{ij}\}_{J \in I}\), where \(I = I_1 \cup I_2\). Let \(\preceq\) be a linear ordering on \(I\), with \(j_1 \preceq j_2\) for every \(j_1 \in I_1\) and \(j_2 \in I_2\). Note that the restrictions of \(\preceq\) to \(I_1\) and \(I_2\) define linear orderings on those sets.

Let \(M(I)\) and \(x_M \in UA\), \(M \in M(I)\), be as in Section 25.10, so that the family of \(x_M\), \(M \in M(I)\), is a basis for \(UA\), as a module over \(k\), by the Poincaré–Birkhoff–Witt theorem. Of course, we get analogous bases for \(UA_1\) and \(UA_2\), as modules over \(k\), indexed by \(M(I_1)\) and \(M(I_2)\), respectively. This leads to a basis for \((UA_1) \otimes (UA_2)\), as a module over \(k\), indexed by \(M(I_1) \times M(I_2)\). Elements of \(M(I_1)\) and \(M(I_2)\) can be combined to get elements of \(M(I)\), and every element of \(M(I)\) can be obtained by combining unique elements of \(M(I_1)\) and \(M(I_2)\) in this way. This defines a one-to-one correspondence between \(M(I_1) \times M(I_2)\) and \(M(I)\).

The linear mapping from \((UA_1) \otimes (UA_2)\) into \(UA\) determined by (25.12.2) defines a one-to-one correspondence between the bases for \((UA_1) \otimes (UA_2)\) and \(UA\), as modules over \(k\), defined in the preceding paragraph. In particular, this linear mapping is an isomorphism between \((UA_1) \otimes (UA_2)\) and \(UA\), as modules over \(k\), as in Corollary 2 on p16 of [24].

In this case, we also have an induced isomorphism from \((\text{gr } UA_1) \otimes (\text{gr } UA_2)\) onto \(\text{gr } UA\), as on p16 of [24]. More precisely, remember that \(\text{gr } UA\) is isomorphic to \(SA\), as in the Poincaré–Birkhoff–Witt theorem, and similarly \(\text{gr } UA_1\), \(\text{gr } UA_2\) are isomorphic to \(SA_1\), \(SA_2\), respectively. The corresponding isomorphism from \((SA_1) \otimes (SA_2)\) onto \(SA\) can be seen in terms of bases here, or as in Section 25.7.

### 25.13 Some related filtrations

Let us look a bit more at the filtrations associated to the type of situation mentioned at the end of the previous section. Let \(k\) be a commutative ring with a multiplicative identity element again, let \((A, [\cdot, \cdot])_A\) be a Lie algebra over \(k\), and let \(A_1\), \(A_2\) be Lie subalgebras of \(A\). Also let \(UA\), \(UA_1\), and \(UA_2\) be universal enveloping algebras of \(A\), \(A_1\), and \(A_2\), respectively, with multiplicative identity elements \(e = e_{UA}\), \(e_1 = e_{UA_1}\), and \(e_2 = e_{UA_2}\), and Lie algebra homomorphisms \(i = i_{UA}\), \(i_1 = i_{UA_1}\), and \(i_2 = i_{UA_2}\) from \(A\), \(A_1\), and \(A_2\) into \(UA\), \(UA_1\), and \(UA_2\), respectively. If \(n\) is a nonnegative integer, then the corresponding submodules \(U_nA\), \(U_nA_1\), and \(U_nA_2\) of \(UA\), \(UA_1\), and \(UA_2\), respectively, can be defined as in Section 25.9.

The restriction \(\phi_j\) of \(i\) to \(A_j\) is a Lie algebra homomorphism from \(A_j\) into \(UA\) for \(j = 1, 2\), which leads to a unique algebra homomorphism \(\psi_j\) from \(UA_j\) into \(UA\) such that \(\psi_j(e_j) = e\) and \(\psi_j \circ i = \phi_j\), as before. Using this, we get a
bilinear mapping from \((UA_1) \times (UA_2)\) into \(UA\) as in (25.12.1), which leads to a unique linear mapping from \((UA_1) \otimes (UA_2)\) into \(UA\) that satisfies (25.12.2). If \(w_1 \in U_{n_1}A_1\) and \(w_2 \in U_{n_2}A_2\) for some \(n_1, n_2 \geq 0\), then it is easy to see that

\[
\psi_1(w_1) \psi_2(w_2) \in U_{n_1+n_2}A.
\]  

Let \(n\) be a nonnegative integer, and let \(((UA_1) \otimes (UA_2))_n\) be the submodule of \((UA_1) \otimes (UA_2)\), as a module over \(k\), generated by elements of the form \(w_1 \otimes w_2\), with \(w_1 \in U_{n_1}A_1\), \(w_2 \in U_{n_2}A_2\), and \(n_1 + n_2 \leq n\). Thus the linear mapping from \((UA_1) \otimes (UA_2)\) into \(UA\) mentioned before maps \(((UA_1) \otimes (UA_2))_n\) into \(U_nA\) for every \(n \geq 0\), by (25.13.1).

Observe that

\[
((UA_1) \otimes (UA_2))_n \subseteq ((UA_1) \otimes (UA_2))_{n+1}
\]

for every \(n \geq 0\), and that \(\bigcup_{n=0}^{\infty}((UA_1) \otimes (UA_2))_n = (UA_1) \otimes (UA_2)\). Remember that \(gr_nUA\) is defined as in Section 25.9 for each nonnegative integer \(n\), and similarly for \(gr_n(UA_1, gr_n(UA_2)\). Put

\[
gr_n((UA_1) \otimes (UA_2)) = ((UA_1) \otimes (UA_2))_n / ((UA_1) \otimes (UA_2))_{n-1}
\]

for each positive integer \(n\), where the quotient is defined as a module over \(k\), and \(gr_0((UA_1) \otimes (UA_2))_0 = ((UA_1) \otimes (UA_2))_0\). Using the linear mapping from \((UA_1) \otimes (UA_2)\) into \(UA\) mentioned in the preceding paragraph, we get an induced linear mapping from \(gr_n((UA_1) \otimes (UA_2))_n\) into \(gr_nUA\) for every \(n \geq 0\). If we put

\[
gr((UA_1) \otimes (UA_2)) = \bigoplus_{n=0}^{\infty} gr_n((UA_1) \otimes (UA_2)),
\]

then we get a homomorphism from \(gr((UA_1) \otimes (UA_2))\) into \(grUA\), as modules over \(k\).

Of course,

\[
(w_1, w_2) \mapsto w_1 \otimes w_2
\]

defines a bilinear mapping from \((UA_1) \times (UA_2)\) into \((UA_1) \otimes (UA_2)\). If \(n_1, n_2\) are nonnegative integers, then this bilinear mapping sends \((U_{n_1}A_1) \times (U_{n_2}A_2)\) into \(((UA_1) \otimes (UA_2))_{n_1+n_2}\). This leads to a bilinear mapping

\[
(gr_{n_1}UA_1) \times (gr_{n_2}UA_2) \rightarrow gr_{n_1+n_2}((UA_1) \otimes (UA_2)).
\]

Using this, we get a bilinear mapping

\[
(grUA_1) \times (grUA_2) \rightarrow gr((UA_1) \otimes (UA_2)).
\]

This leads to a linear mapping

\[
gr((UA_1) \otimes (UA_2)) \rightarrow gr((UA_1) \otimes (UA_2)).
\]
More precisely, if \( n \) is a nonnegative integer, then then \( \text{gr}_{n}(\mathfrak{u}(\mathfrak{u}(A_{1}) \otimes \mathfrak{u}(A_{2})) \) is generated, as a module over \( k \), by the images of \((\text{gr}_{n_{1}} \mathfrak{u}(A_{1}) \times (\text{gr}_{n_{2}} \mathfrak{u}(A_{2})\), with \( n_{1} + n_{2} = n \). This means that

\[ (\text{gr}(\mathfrak{u}(A_{1})) \otimes (\text{gr}(\mathfrak{u}(A_{2})) \text{ maps onto } \text{gr}(\mathfrak{u}(A_{1}) \otimes (\text{gr}(A_{2}))), \]

under the mapping just mentioned.

Suppose that \( A_{1} \) and \( A_{2} \) are free as modules over \( k \), with bases \( \{x_{j}\}_{j \in I_{1}} \) and \( \{x_{j}\}_{j \in I_{2}} \), respectively. Suppose that \( I_{1} \) and \( I_{2} \) are also linearly ordered, so that \( \mathcal{M}(I_{1}), \mathcal{M}(I_{2}) \) can be defined as in Section 25.10. Using the Poincaré–Birkhoff–Witt theorem, we get bases for \( \mathfrak{u}(A_{1}) \) and \( \mathfrak{u}(A_{2}) \), as modules over \( k \), indexed by \( \mathcal{M}(I_{1}) \) and \( \mathcal{M}(I_{2}) \), respectively. In particular, the mapping indicated in (25.13.8) is an isomorphism between modules over \( k \) in this situation. Using these bases for \( \mathfrak{u}(A_{1}) \) and \( \mathfrak{u}(A_{2}) \), we get a basis for \( \mathfrak{u}(A_{1}) \otimes \mathfrak{u}(A_{2}) \) as well, as a module over \( k \).

Suppose that \( A \) corresponds to the direct sum of \( A_{1} \) and \( A_{2} \), as a module over \( k \). Thus \( A \) is free as a module over \( k \) too, with basis \( \{x_{j}\}_{j \in I} \), where \( I = I_{1} \cup I_{2} \), and \( I_{1} \cap I_{2} = \emptyset \). Let \( \preceq \) be a linear ordering on \( I \), with \( j_{1} \preceq j_{2} \) for every \( j_{1} \in I_{1} \) and \( j_{2} \in I_{2} \), and let us take \( I_{1} \) and \( I_{2} \) to be linearly ordered by the restrictions of \( \preceq \) to the subsets of \( I \). Using the Poincaré–Birkhoff–Witt theorem again, we get a basis for \( \mathfrak{u}(A) \), as a module over \( k \), indexed by the corresponding set \( \mathcal{M}(I) \). As in the previous section, the linear mapping from \( \mathfrak{u}(A_{1}) \otimes \mathfrak{u}(A_{2}) \) into \( \mathfrak{u}(A) \) determined by (25.12.2) is an isomorphism between these spaces, as modules over \( k \), because it defines a one-to-one correspondence between their bases.

Similarly, the homomorphism from \( \text{gr}(\mathfrak{u}(A_{1}) \otimes \mathfrak{u}(A_{2})) \) into \( \text{gr}(\mathfrak{u}(A)) \), as modules over \( k \), mentioned earlier is an isomorphism under these conditions. This can be composed with the isomorphism indicated in (25.13.8), to get an isomorphism from \( \text{gr}(\mathfrak{u}(A_{1}) \otimes \mathfrak{u}(A_{2})) \) onto \( \text{gr}(\mathfrak{u}(A)) \), as modules over \( k \), under these conditions.

### 25.14 The diagonal mapping

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( (A, [\cdot, \cdot], \cdot) \) be a Lie algebra over \( k \). Also let \( \mathfrak{u}(A) \) be a universal enveloping algebra of \( A \), with multiplicative identity element \( e = e_{\mathfrak{u}(A)} \) and Lie algebra homomorphism \( i = i_{\mathfrak{u}(A)} \) from \( A \) into \( \mathfrak{u}(A) \), as before. Of course, \( A \times A \) may be considered as a Lie algebra over \( k \) too, where addition, scalar multiplication, and the Lie bracket are defined coordinatewise. Note that \( A_{1} = A \times \{0\} \) and \( A_{2} = \{0\} \times A \) are Lie subalgebras of \( A \times A \), and that \( A \times A \) corresponds to the direct sum of \( A_{1} \) and \( A_{2} \), as a Lie algebra over \( k \). We may consider \( \mathfrak{u}(A) \) to be a universal enveloping algebra of \( A_{1} \) and \( A_{2} \), using the obvious isomorphisms \( a \mapsto (a, 0) \) and \( a \mapsto (0, a) \) from \( A \) onto \( A_{1} \) and \( A_{2} \), respectively.

If \( a_{1}, a_{2} \in A \), then put

\[ (25.14.1) \quad \rho((a_{1}, a_{2})) = i(a_{1}) \otimes e + e \otimes i(a_{2}), \]
25.14. THE DIAGONAL MAPPING

which defines a mapping from $A \times A$ into $(UA) \otimes (UA)$. More precisely, $\rho$ is a Lie algebra homomorphism from $A \times A$ into $(UA) \otimes (UA)$, as in Section 25.7. This leads to an algebra homomorphism $\theta$ from a universal enveloping algebra of $A \times A$ into $(UA) \otimes (UA)$, which is an algebra isomorphism in this situation, as before. One can use this to consider $(UA) \otimes (UA)$ as a universal enveloping algebra of $A \times A$.

Clearly $a \mapsto (a, a)$ is a Lie algebra homomorphism from $A$ into $A \times A$. We can compose this with $\rho$ to get a Lie algebra homomorphism

$$a \mapsto i(a) \otimes e + e \otimes i(a)$$

from $A$ into $(UA) \otimes (UA)$. This leads to a unique algebra homomorphism $\Delta$ from $UA$ into $(UA) \otimes (UA)$, where $\Delta(e) = e \otimes e$ and $\Delta \circ i$ is the same as (25.14.2). This is called the diagonal map, as in Definition 5.1 and Proposition 5.2 on p16 of [24].

An element $\alpha$ of $UA$ is said to be primitive if

$$\Delta(\alpha) = \alpha \otimes e + e \otimes \alpha,$$

as in Definition 5.3 on p17 of [24]. Every element of $i(A)$ is primitive, as in the preceding paragraph.

If $t \in k$ and $n \in \mathbb{Z}_+$, then $n \cdot t$ denotes the sum of $n$ $t$'s in $k$, as usual. Suppose that $k$ is torsion-free, so that $n \cdot t = 0$ implies $t = 0$. If $A$ is free as a module over $k$, then Theorem 5.4 on p17 of [24] states that the only primitive elements of $UA$ are in $i(A)$.

Suppose first that $A$ is commutative as a Lie algebra over $k$, as in [24]. Let $\{x_j\}_{j \in I}$ be a basis for $A$, as a module over $k$, and let $X_j$, $j \in I$, be a commuting family of indeterminates. There is a unique module homomorphism from $A$ into the polynomial algebra $k[[X_j : j \in I]]$ that sends $x_j$ to $X_j$ for every $j \in I$. Using this, $k[[X_j : j \in I]]$ may be considered as a symmetric algebra of $A$, as in Section 25.8. This means that $k[[X_j : j \in I]]$ is a universal enveloping algebra of $A$ in this case.

Let $X'_j$ and $X''_j$ be additional indeterminates for each $j \in I$, which commute with each other and with $X'_j$, $X''_j$, $l \in I$. The tensor product of $k[[X_j : j \in I]]$ with itself can be identified with the polynomial algebra

$$k[[X'_j, X''_j : j \in I]],$$

where $X'_j$ corresponds to $X_j \otimes 1$ and $X''_j$ corresponds to $1 \otimes X_j$ for every $j \in I$. In this situation, $\Delta$ corresponds to the mapping from $k[[X_j : j \in I]]$ to (25.14.4) that sends a formal polynomial $f(X)$ in the $X_j$'s to the formal polynomial $f(X' + X'')$ in the $X'_j$'s and $X''_j$'s, which replaces $X_j$ with $X'_j + X''_j$ for every $j \in I$. The primitive elements correspond to $f(X) \in k[[X_j : j \in I]]$ such that

$$f(X' + X'') = f(X') + f(X'').$$

If $f(X)$ is homogeneous of degree $n$ for some nonnegative integer $n$, then this would imply that

$$2^n \cdot f(X) = f(X + X) = f(X) + f(X) = 2 \cdot f(X),$$
which is to say that $(2^n - 2) \cdot f(X) = 0$. If $n \neq 1$, then this implies that $f(X) = 0$, because $k$ is torsion-free, by hypothesis. Otherwise, one can apply this argument to the homogeneous components of $f(X)$, which satisfy the same condition.

Let us now consider the case where $A$ is not necessarily commutative as a Lie algebra over $k$, as in [24] again. Remember that $a \mapsto (a,a)$ is a Lie algebra homomorphism from $A$ into $A \times A$. If $U(A \times A)$ is a universal enveloping algebra of $A \times A$, then we can compose the previous Lie algebra homomorphism from $A$ into $A \times A$ with the usual Lie algebra homomorphism from $A \times A$ into $U(A \times A)$. This leads to a unique algebra homomorphism from $UA$ into $U(A \times A)$, with the usual properties. It is easy to see that $\Delta$ is the same as the composition of this homomorphism with the algebra isomorphism from $U(UA)$ onto $(UA) \otimes (UA)$ mentioned earlier.

The algebra homomorphism from $UA$ into $U(A \times A)$ mentioned in the preceding paragraph induces an algebra homomorphism from $\text{gr} UA$ into $\text{gr} U(A \times A)$. We also have that $\theta$ induces an algebra isomorphism from $\text{gr} U(A \times A)$ onto $(\text{gr} UA) \otimes (\text{gr} UA)$, as in the previous sections. Using this, $\Delta$ induces an algebra homomorphism $\text{gr} \Delta$ from $\text{gr} UA$ into $(\text{gr} UA) \otimes (\text{gr} UA)$. Remember that $\text{gr} UA$ is isomorphic to $SA$, by the Poincaré-Birkhoff-Witt theorem. Thus $\text{gr} \Delta$ corresponds to an algebra homomorphism from $SA$ into $(SA) \otimes (SA)$.

Of course, $SA$ is the same as a universal enveloping algebra of $A$ as a commutative Lie algebra over $k$. The algebra homomorphism from $SA$ into $(SA) \otimes (SA)$ mentioned in the preceding paragraph corresponds exactly to the one considered in the first case. This can be seen by considering elements of $\text{gr}_1 UA$ that come from elements of $A$ in the usual way.

Let $x$ be a primitive element of $UA$, and suppose that $x \in U_n A$ for some $n \geq 0$. This implies the analogous condition for the image of $x$ in $\text{gr} U(A \times A)$ with respect to $\text{gr} \Delta$. If $n > 1$, then it follows that the image of $x$ in $\text{gr}_n UA$ is equal to 0, as in the first case. This means that $x \in U_{n-1} A$, and we can repeat the process to get that $x \in U_1 A$.

Thus $x = te + i(a)$ for some $t \in k$ and $a \in A$, which implies that

\begin{equation}
\Delta(x) = t(e \otimes e) + i(a) \otimes e + e \otimes i(a).
\end{equation}

Observe that

\begin{equation}
x \otimes e + e \otimes x = t(e \otimes e) + i(a) \otimes e + t(e \otimes e) + e \otimes i(a).
\end{equation}

The hypothesis that $x$ be primitive means that (25.14.7) and (25.14.8) are the same, so that $t = 0$. This shows that $x = i(a) \in i(A)$, as desired.

### 25.15 More on primitive elements

Let us return to the situation considered at the beginning of the previous section. Thus we let $k$ be a commutative ring with a multiplicative identity element again, and $(A, [\cdot, \cdot])$ be a Lie algebra over $k$. Also let $UA$ be a universal
enveloping algebra of $A$, with multiplicative identity element $e = e_{UA}$ and Lie algebra homomorphism $i = i_{UA}$ from $A$ into $UA$. Remember that the diagonal map $\Delta$ is an algebra homomorphism from $UA$ into $(UA) \otimes (UA)$.

Let $PUA$ be the set of primitive elements of $UA$. It is easy to see that $PUA$ is a submodule of $UA$, as a module over $k$. Exercise 1 on p17 of [24] asks one to show that if $x, y \in PUA$, then $[x, y] = x y - y x \in PUA$. This means that

\begin{equation}
\Delta([x, y]) = ([x, y]) \otimes e + e \otimes ([x, y]).
\end{equation}

Of course, $\Delta([x, y]) = [\Delta(x), \Delta(y)]$, because $\Delta$ is an algebra homomorphism. It follows that

\begin{equation}
\Delta([x, y]) = [x \otimes e + e \otimes x, y \otimes e + e \otimes y],
\end{equation}

because $x, y$ are primitive by hypothesis. It is easy to obtain (25.15.1) from (25.15.2), because of the way multiplication is defined in $(UA) \otimes (UA)$, as in Section 25.2.

Of course, we may as well suppose that $k \neq \{0\}$, so that $1 \neq 0$ in $k$. Let us suppose from now on in this section that there is a prime number $p$ such that $p \cdot 1 = 0$ in $k$, as in Exercise 2 on p17 of [24]. Of course, this means that $p \cdot t = 0$ for every $t \in k$.

Let $n$ be a positive integer that is not a multiple of $p$, so that there is a positive integer $m$ such that $m n - 1$ is a multiple of $p$. If $t \in k$, then

\begin{equation}
m \cdot (n \cdot t) = (m n) \cdot t = t.
\end{equation}

In particular, if $n \cdot t = 0$, then $t = 0$.

If $y \in PUA$, then part (a) of Exercise 2 on p17 of [24] asks us to show that $y^p \in PUA$.

Observe that

\begin{equation}
\Delta(y^p) = (\Delta(y))^p = (y \otimes e + e \otimes y)^p,
\end{equation}

because $\Delta$ is an algebra homomorphism. The right side can be expanded into a sum, using the binomial theorem, because $y \otimes e$ commutes with $e \otimes y$ in $(UA) \otimes (UA)$. It is well known and easy to see that the binomial coefficient $\binom{p}{j}$ is a multiple of $p$ when $j = 1, \ldots, p - 1$. It follows that

\begin{equation}
\Delta(y^p) = (y \otimes e)^p + (e \otimes y)^p = y^p \otimes e + e \otimes y^p,
\end{equation}

as desired.

Let us suppose for the rest of the section that $A$ is free as a module over $k$, and let $\{x_j\}_{j \in I}$ be a basis for $A$, as a module over $k$. If $j \in I$, then $i(x_j) \in PUA$, as in the previous section. This implies that

\begin{equation}
i(x_j)^\nu \in PUA
\end{equation}

for every positive integer $\nu$, by (25.15.4). Part (b) of Exercise 2 on p17 of [24] asks us to show that the family of elements of $PUA$ of the form

\begin{equation}
i(x_j)^\nu, \text{ with } j \in I \text{ and } \nu \in \mathbb{Z}_+ \cup \{0\},
\end{equation}
is a basis for $PUA$, as a module over $k$. It suffices to show that every element of $PUA$ can be expressed as a finite linear combination of these elements, because of the Poincaré–Birkhoff–Witt theorem.

Suppose first that $A$ is commutative as a Lie algebra over $k$, as in the previous section, and let $X_j$, $j \in I$, be a commuting family of indeterminates again. As before, the polynomial algebra $k[[X_j : j \in I]]$ may be considered as a symmetric algebra of $A$, using the unique module homomorphism from $A$ into this algebra that sends $x_j$ to $X_j$ for every $j \in I$. This is the same as a universal enveloping algebra of $A$ in this case. Let $X'_j$ and $X''_j$ be additional indeterminates for each $j \in I$ again, which commute with each other and $X'_j$, $X''_j$, $l \in I$, so that the tensor product of $k[[X_j : j \in I]]$ with itself can be identified with

$$
(25.15.9) \quad k[[X'_j, X''_j : j \in I]],
$$

as before. Remember that $\Delta$ corresponds in this situation to the mapping from $k[[X_j : j \in I]]$ to (25.15.9) that sends a formal polynomial $f(X)$ in the $X_j$’s to the formal polynomial $f(X' + X'')$ in the $X'_j$ and $X''_j$’s. The primitive elements correspond to the formal polynomials $f(X)$ such that

$$
(25.15.10) \quad f(X' + X'') = f(X') + f(X''),
$$

as before. In particular, this implies that the constant term in $f(X)$ is equal to 0.

If $f(X)$ satisfies (25.14.5, 2), then we would like to show that $f(X)$ can be expressed as a finite linear combination of the monomials $X'_j$, with $j \in I$ and $\nu \in \mathbb{Z}_+ \cup \{0\}$. We may as well suppose that the linear terms in $f(X)$ are equal to 0, which would correspond to $\nu = 0$. Under these conditions, one can check that the formal partial derivative of $f(X)$ in the $j$th variable is equal to 0 for every $j \in I$. To see this, consider the sum of the terms in $f(X' + X'')$ with exactly one factor of $X''_j$ and no factors of $X''_j$ when $j \neq l$, and any number of factors of $X'_j$ for $l \in I$. This sum is equal to the formal partial derivative of $f(X')$ in the $j$th variable times $X''_j$. This would have to be a scalar multiple of $X''_j$, by (25.15.10), which would be a term in $f(X'')$. In fact, this has to be equal to 0, because the linear terms in $f(X)$ are supposed to be equal to 0.

If $j \in I$, then $f(X)$ can be identified with a formal polynomial in $X_j$ whose coefficients are formal polynomials in $X_l$, $j \neq l$. The formal partial derivative of $f(X)$ in the $j$th variable corresponds to the usual formal derivative of $f(X)$ as a formal polynomial in this way. In this situation, the condition that the derivative be equal to 0 implies that this formal polynomial in $X_j$ can be expressed as a formal polynomial in $X'_j$. This means that $f(X)$ can be expressed as a formal polynomial in $X'_j$, $j \in I$, with coefficients in $k$. Let us express this formally as

$$
(25.15.11) \quad f(X) = g(X').
$$

Using the binomial theorem again, we get that $(X'_j + X''_j)^p = (X'_j)^p + (X''_j)^p$ for every $j \in I$. This implies that $g$ corresponds to a primitive element in the polynomial algebra as well. Note that the linear terms in $g$ correspond to terms
in \( f(X) \) of the form \( X_j^p, j \in I \). We can subtract the linear terms from \( g \), and repeat the previous argument. After a finite number of steps, we can express \( f(X) \) as a linear combination of monomials \( X_j^p \), with \( j \in I, \nu \geq 0 \), as desired.

If \( A \) is not necessarily commutative as a Lie algebra over \( k \), then one can reduce to the previous case, as in the preceding section. Let \( x \in PUA \) be given, and suppose that \( x \in U_nA \) for some \( n \geq 0 \). This implies that the image of \( x \) in \( \text{gr}_nUA \) satisfies the analogous primitivity condition with respect to \( \text{gr} \Delta \), as before. Using the previous case, we can express the image of \( x \) in \( \text{gr}_nUA \) in the analogous way. This means that \( x \) can be expressed as the sum of a finite linear combination of elements of \( PUA \) of the form (25.15.8) and an element of \( U_{n-1}A \). This element of \( U_{n-1}A \) is primitive as well, because \( PUA \) is a submodule of \( UA \), as a module over \( k \). Thus we can repeat the process, to get that \( x \) is a finite linear combination of elements of \( PUA \) of the form (25.15.8), as desired.

If \( x, y \in i(A) \), then part (c) of Exercise 2 on p17 of [24] asks us to show that

\[
(x + y)^p - x^p - y^p \in i(A).
\]

(25.15.12)

Note that \( x^p, y^p, (x + y)^p \in PUA \), as in (25.15.7), and because \( i(A) \subseteq PUA \). This implies that

\[
(x + y)^p - x^p - y^p \in PUA,
\]

(25.15.13)

because \( PAU \) is a submodule of \( UA \), as a module over \( k \).

We also have that

\[
(x + y)^p - x^p - y^p \in U_{p-1}A.
\]

(25.15.14)

More precisely, \( (x + y)^p - x^p - y^p \in U_pA \), and (25.15.14) is the same as saying that its image in \( \text{gr}_pUA \) is equal to 0. This can be obtained from the binomial theorem, because \( \binom{p}{l} \) is a multiple of \( p \) when \( l = 1, \ldots, p - 1 \). It is easy to see that (25.15.12) follows from (25.15.13) and (25.15.14), using the basis (25.15.8) for \( PUA \).
Chapter 26

Free Lie algebras

26.1 Magmas and their homomorphisms

A magma is a set $M$ together with a mapping from $M \times M$ into $M$, which may be expressed as

\[(26.1.1) \quad (x, y) \mapsto xy.\]

This is Definition 1.1 on p18 of [24]. A mapping $\phi$ from $M$ into another magma $N$ is said to be a magma homomorphism if

\[(26.1.2) \quad \phi(xy) = \phi(x) \phi(y)\]

for every $x, y \in M$.

Let $X$ be a set, and put $X_1 = X$. If $n \geq 2$ is an integer, then $X_n$ can be defined recursively as the disjoint union of $X_j \times X_l$, where $j, l \in \mathbb{Z}_+$ and $j + l = n$. Let $M_X$ be the disjoint union of the $X_n$'s, $n \in \mathbb{Z}_+$. There is a natural binary operation on $M_X$, defined using the embedding of $X_j \times X_l$ into $X_{j+l}$ for every $j, l \geq 1$, by construction. This makes $M_X$ into a magma, which is the free magma on $X$, as on p18 of [24]. The elements of $M_X$ may be described as non-associative words on $X$. If $n \in \mathbb{Z}_+$, then the elements of $X_n$ are said to have length $n$.

If $f$ is any mapping from $X$ into a magma $N$, then there is a unique extension of $f$ to a magma homomorphism $F$ from $M_X$ into $N$. This is Theorem 1.2 on p18 of [24]. More precisely, if $j, l \in \mathbb{Z}_+$, $x \in X_j$, and $y \in X_l$, then $(x, y) \in X_{j+l}$, and $F$ should satisfy

\[(26.1.3) \quad F((x, y)) = F(x) F(y).\]

This can be used to define $F$ recursively on $M_X$.

Of course, $M_X$ is generated by elements of $X$ in the obvious way. Every element of $M_X \setminus X$ can be expressed in a unique way as a product of elements of $M_X$, as mentioned on p18 of [24].
26.2 Magmas and algebras

Let $k$ be a commutative (associative) ring with a multiplicative identity element, and let $M$ be a nonempty magma. We can define an algebra $A$ over $k$ in the strict sense using $M$ in the usual way, whose elements are formal sums

$\sum_{m \in M} c_m m$

with $c_m \in k$ for every $m \in M$, and $c_m = 0$ for all but finitely many $m \in M$. More precisely, this can be defined as the space of $k$-valued functions on $M$ with finite support, and it is convenient to express the elements of $A$ as formal sums in this way. Of course, $A$ is a module over $k$ with respect to pointwise addition and scalar multiplication of functions on $M$, which corresponds to termwise addition and scalar multiplication of these formal sums. Multiplication on $M$ extends to a bilinear mapping from $A \times A$ into $A$, which defines multiplication on $A$.

Let $X$ be a nonempty set, and let $M_X$ be the free magma on $X$, as in the previous section. Using $M_X$, we get an algebra $A_X(k) = A_M$ over $k$ in the strict sense, as in the preceding paragraph. This is called the free algebra on $X$ over $k$, as in Definition 2.1 on p18 of [24].

Let $B$ be another algebra over $k$ in the strict sense, and let $f$ be a mapping from $X$ into $B$. Under these conditions, there is a unique extension of $f$ to an algebra homomorphism $F$ from $A_X$ into $B$, as in Theorem 2.2 on p19 of [24]. Note that $B$ may be considered as a magma with respect to multiplication. Thus $f$ can first be extended to a magma homomorphism from $M_X$ into $B$, as in the previous section. This can be extended to a mapping $F$ from $A_X$ into $B$ that is linear over $k$. It is easy to see that $F$ is an algebra homomorphism on $A_X$, because it is a magma homomorphism on $M_X$. The uniqueness of $F$ follows from the fact that $A_X$ is generated by $X$, as an algebra over $k$.

Remember that the length of an element of $M_X$ is defined as in the previous section. Let us say that an element

$a = \sum_{m \in M_X} c_m m$

of $A_X$ is homogeneous of degree $n$ for some positive integer $n$ if $c_m = 0$ for every $m \in M_X$ such that the length of $m$ is not equal to $n$. The set $A_X^n(k) = A_X^n$ of elements of $A_X$ that are homogeneous of degree $n$ is a submodule of $A_X$, as a module over $k$, and $A_X$ corresponds to the direct sum of these submodules, as a module over $k$. If $a_1, a_2 \in A_X$ are homogeneous of degrees $n_1, n_2 \geq 1$, respectively, then $a_1 a_2$ is homogeneous of degree $n_1 + n_2$. Thus $A_X$ is graded as an algebra over $k$ in the strict sense in this way, as mentioned on p19 of [24].

26.3 Constructing free Lie algebras

Let $k$ be a commutative ring with a multiplicative identity element again, and let $X$ be a nonempty set. Also let $A_X$ be the free algebra on $X$ over $k$, as in the
previous section. Consider the two-sided ideal $I_X$ of $A_X$ generated by elements of the form

\begin{equation}
\tag{26.3.1}
a a,
\end{equation}

with $a \in A_X$, and

\begin{equation}
\tag{26.3.2}
J(a, b, c) = (ab)c + (bc)a + (ca)b,
\end{equation}

with $a, b, c \in A_X$. Thus an element of $I_X$ is a finite sum of elements of $A_X$, each of which can be obtained by starting with an element of the form (26.3.1) or (26.3.2), and multiplying it on the left and right by elements of $A_X$. More precisely, one can multiply elements of the form (26.3.1) or (26.3.2) by elements of $A_X$ on the left or right by elements of $A_X$ repeatedly, and in any order.

Thus the quotient

\begin{equation}
\tag{26.3.3}
A_X/I_X
\end{equation}

is defined as an algebra over $k$ in the strict sense. It is easy to see that (26.3.3) is a Lie algebra over $k$, because elements of $A_X$ of the form (26.3.1) and (26.3.2) are mapped to 0 in (26.3.3). More precisely, the Lie bracket on (26.3.3) is the binary operation that corresponds to multiplication on $A_X$. This is called the free Lie algebra on $X$ over $k$, as in Definition 3.1 on p19 of [24]. This may be denoted $L_X(k)$, or $L_X$ when the choice of $k$ is clear.

Of course, there is a natural mapping from $X$ into $L_X$, obtained by restricting the quotient mapping from $A_X$ onto $L_X$ to $X$, as a subset of $A_X$. Note that $L_X$ is generated as an algebra over $k$ by the image of $X$ in $L_X$, because $A_X$ is generated as an algebra over $k$ by $X$, by construction.

Let $B$ be a Lie algebra over $k$, and let $f$ be a mapping from $X$ into $B$. As in the previous section, there is a unique extension of $f$ to an algebra homomorphism $F$ from $A_X$ into $B$. This extension maps elements of $A_X$ of the form (26.3.1) and (26.3.2) to 0 in $B$, because $B$ is a Lie algebra over $k$. This implies that $I_X$ is contained in the kernel of $F$. This leads to an algebra homomorphism from $L_X$ into $B$. By construction, the composition of this algebra homomorphism with the natural mapping from $X$ into $L_X$ is $f$. This algebra homomorphism from $L_X$ into $B$ is uniquely determined by this property, because $L_X$ is generated as an algebra by the image of $X$, as in the preceding paragraph.

This property is essentially used to define free Lie algebras on $X$ on p94f of [13]. It is easy to see that the free Lie algebra on $X$ is uniquely determined by this property, up to a unique isomorphism, as on p95 of [13]. The existence of a free Lie algebra on $X$ is obtained another way on p95 of [13], and we shall return to this later.

Let $Y$ be another nonempty set, and let $L_Y$ be the corresponding free algebra on $Y$ over $k$. If $f$ is a mapping from $X$ into $Y$, then we get a mapping from $X$ into $L_Y$, by composing $f$ with the natural mapping from $Y$ into $L_Y$. This leads to a unique Lie algebra homomorphism from $L_X$ into $L_Y$, whose composition with the natural mapping from $X$ into $L_X$ is the mapping from $X$ into $L_Y$ just mentioned, as before. This is one of the properties of free Lie algebras mentioned on p19 of [24].
26.4. Degree 2 in $L_X$

Remember that $A_X$ is graded as an algebra over $k$ in the strict sense, as in the previous section. Let $\mathcal{I}_X^\#$ be the collection of elements of $A_X$ whose homogeneous components are elements of $\mathcal{I}_X$. Thus $\mathcal{I}_X^\# \subseteq \mathcal{I}_X$ automatically, and we would like to verify that

\[ \mathcal{I}_X^\# = \mathcal{I}_X \]

as on p19 of [24]. It suffices to check that $\mathcal{I}_X$ is generated by homogeneous elements of $A_X$, as a two-sided ideal in $A_X$.

Let $a \in A_X$ be given, and let $a_n$ be the homogeneous component of $a_n$ for each positive integer $n$, so that $a_n = 0$ for all but finitely many $n \geq 1$, and $a = \sum_{n=1}^\infty a_n$. Observe that

\[ a \cdot a = \sum_{n=1}^\infty a_n a_n + \sum_{n=2}^\infty \sum_{i=1}^{n-1} (a_i a_n + a_n a_i). \]

Clearly $a_n a_n \in \mathcal{I}_X$ for every $n \geq 1$, and

\[ a_i a_n + a_n a_i = (a_i + a_n) (a_i + a_n) - a_i a_i - a_n a_n \]

is an element of $\mathcal{I}_X$ when $l < n$. This implies that $a \cdot a$ can be expressed as a finite sum of homogeneous elements of $\mathcal{I}_X$.

Similarly, let $b, c \in A_X$ be given, with homogeneous components $b_n$, $c_n$ for every $n \geq 1$, respectively. It is easy to see that

\[ J(a, b, c) = \sum_{j=1}^\infty \sum_{l=1}^\infty \sum_{n=1}^\infty J(a_j, b_l, c_n). \]

Thus $J(a, b, c)$ can be expressed as a finite sum of homogeneous elements of $\mathcal{I}_X$ too. It follows that (26.3.4) holds, as before.

Let us say that an element of $L_X$ is homogeneous of degree $n \in \mathbb{Z}_+$ if it is the image of an element of $A_X$ that is homogeneous of degree $n$ under the natural quotient mapping. This makes $L_X$ a graded algebra over $k$ as well, because of (26.3.4), as on p19 of [24].

In particular, 0 is the only element of $\mathcal{I}_X$ that is homogeneous of degree 1. Thus the quotient mapping from $A_X$ onto $L_X$ is injective on the submodule of $A_X$ consisting of elements that are homogeneous of degree 1. This corresponds to part of the statement at the top of p20 of [24].

26.4 Degree 2 in $L_X$

Let $k$ be a commutative ring with a multiplicative identity element, and let $A$ be a module over $k$. Thus $A \otimes A$ can be defined as a module over $k$ in the usual way. Let $B$ be the submodule of $A \otimes A$ generated by elements of $A \otimes A$ of the form $a \otimes a$, with $a \in A$. Equivalently, $B$ consists of finite linear combinations of elements of this form, with coefficients in $k$. In particular, if $x, y \in A$, then

\[ x \otimes y + y \otimes x = (x + y) \otimes (x + y) - x \otimes x - y \otimes y \]
is an element of $B$. We can define $\bigwedge^2 A$ to be the quotient
\[(26.4.2) \quad (A \otimes A)/B,\]
as a module over $k$. If $x, y \in A$, then let $x \wedge y$ be the image of $x \otimes y$ under the quotient mapping from $A \otimes A$ onto (26.4.2). By construction, $a \wedge a = 0$ for every $a \in A$, and $x \wedge y = -y \wedge x$ for every $x, y \in A$.

Consider $A \times (\bigwedge^2 A)$, as a module over $k$ with respect to coordinatewise addition and scalar multiplication. Of course, this is the same as the direct sum of $A$ and $\bigwedge^2 A$, as a module over $k$. If $x, y \in A$ and $u, v \in \bigwedge^2 A$, then put
\[(26.4.3) \quad [(x, u), (y, v)] = (0, x \wedge y).\]

One can check that $A \times (\bigwedge^2 A)$ is a Lie algebra over $k$ with respect to this definition of Lie bracket, as on p2 of [24].

Let $X$ be a nonempty set, and let $A_X$ be as in Section 26.2. Remember that $A_X$ is a graded as an algebra over $k$ in the strict sense, and let $A$ be the submodule of $A$ consisting of elements that are homogeneous of degree 1. The elements of $A$ correspond to formal sums of elements of $X$ with coefficients in $k$, where all but finitely many of the coefficients are equal to 0. This can be identified with the space of all $k$-valued functions on $X$ with finite support, and is free as a module over $k$.

Similarly, the elements of $A_X$ that are homogeneous of degree 2 correspond to formal sums of elements of $X \times X$ with coefficients in $k$, where all but finitely many of the coefficients are equal to 0. The submodule of $A_X$ consisting of elements that are homogeneous of degree 2 is free as a module over $k$, and can be identified with $A \otimes A$. The intersection of this submodule with the ideal $I_X$ in $A_X$ defined in the previous section corresponds in this way exactly to the submodule $B$ of $A \otimes A$ mentioned earlier. Remember that the free Lie algebra $L_X$ on $X$ over $k$ is graded as an algebra over $k$, as in the previous section. The submodule of $L_X$ consisting of homogeneous elements of degree 2 corresponds to $\bigwedge^2 A$, as on p20 of [24].

Of course, we can identify $A$ and $\bigwedge^2 A$ with submodules of $A \times (\bigwedge^2 A)$, as a module over $k$, in the obvious way. Thus the natural mapping from $X$ into $A$ can be identified with a mapping from $X$ into $A \times (\bigwedge^2 A)$. This leads to a Lie algebra homomorphism from $L_X$ into $A \times (\bigwedge^2 A)$, as in the previous section. This homomorphism sends homogeneous elements of $L_X$ of degree 1 and 2 to their counterparts in $A$ and $\bigwedge^2 A$, as on p20 of [24]. Homogeneous elements of $L_X$ of larger degree are sent to 0 by this homomorphism.

If $a_1, a_2 \in A$, then $a_1, a_2$ can be identified with homogeneous elements of $L_X$ of degree 1, whose Lie bracket $[a_1, a_2]$ is homogeneous of degree 2 in $L_X$. This corresponds to $a_1 \wedge a_2$ in $\bigwedge^2 A$, as before. In particular, if $x, y \in X$, then $x, y$ can be identified with elements of $A$, and $[x, y]$ defined as a homogeneous element of $L_X$ of degree 2.

Let $\preceq$ be a linear ordering on $X$. The collection of elements of $L_X$ of the form $[x, y]$, where $x, y \in X$, $x \preceq y$, and $x \neq y$, is a basis for the set of homogeneous elements of $L_X$ of degree 2, as a module over $k$, as on p20 of [24].
26.5 Free associative algebras

Let $k$ be a commutative ring with a multiplicative identity element, and let $X$ be a nonempty set. Also let $E = E_X$ be the free module over $k$ with the elements of $X$ as a basis, so that the elements of $E$ can be given as formal sums

$$
\sum_{x \in X} c_x x,
$$

with $c_x \in k$ for every $x \in X$, and $c_x = 0$ for all but finitely many $x \in X$. This can be defined as the space of $k$-valued functions on $X$ with finite support, which can be expressed as formal sums in this way for convenience. This is a module over $k$ with respect to termwise addition and scalar multiplication of formal sums, which corresponds to pointwise addition and scalar multiplication of functions on $X$.

The free associative algebra on $X$ over $k$, $\text{Assoc}_X(k) = \text{Assoc}_X$, is defined to be the tensor algebra $TE$ on $E$, as in Definition 4.1 on p20 of [24]. If $f$ is any mapping from $X$ into a module $A$ over $k$, then $f$ has a natural unique extension to a module homomorphism $\phi$ from $E$ into $A$. If $A$ is an associative algebra over $k$ with a multiplicative identity element $e$, then there is a unique algebra homomorphism $\psi$ from $TE$ into $A$ that is equal to $\phi$ on $E = T^1E$ and sends $1 \in k = T^0E$ to $e$, as in Section 25.3.

Of course, $\text{Assoc}_X$ may be considered as a Lie algebra over $k$, using the commutator bracket corresponding to multiplication on $\text{Assoc}_X$. The natural inclusion mapping from $X$ into $E = T^1E$ may be considered as a mapping from $X$ into $\text{Assoc}_X = TE$, which leads to a unique Lie algebra homomorphism $\phi$ from the free Lie algebra $L_X$ on $X$ over $k$ into $\text{Assoc}_X$, as in Section 26.3. More precisely,

$$
\text{the composition of } \phi \text{ with natural mapping from } X \text{ into } L_X
$$

is the same as the natural mapping from $X$ into $\text{Assoc}_X$.

Let $UL_X$ be a universal enveloping algebra of $L_X$, with corresponding multiplicative identity element $e_{UL_X}$ and mapping $i_{UL_X}$ from $L_X$ into $UL_X$. The Lie algebra homomorphism $\phi$ from $L_X$ into $\text{Assoc}_X$ mentioned in the preceding paragraph leads to a unique algebra homomorphism $\Phi$ from $UL_X$ into $\text{Assoc}_X$ such that $\Phi(e_{UL_X}) = 1 \in k = T^0E$ and

$$
\Phi \circ i_{UL_X} = \phi,
$$

as in Section 25.4.

Similarly,

$$
\text{the composition of } \psi \text{ with the natural mapping from } X
$$

is obtained by composing the natural mapping from $X$ into $L_X$ with $i_{UL_X}$.

This leads to a unique algebra homomorphism $\Psi$ from $\text{Assoc}_X$ into $UL_X$, such that

$$
\text{the composition of } \Psi \text{ with the natural mapping from } X
$$
into $\text{Assoc}_X$ is the same as the mapping in (26.5.4),
and which sends $1 \in k = T^0E$ to $e_{UL_X}$.

Thus $\Phi \circ \Psi$ is an algebra homomorphism from $\text{Assoc}_X$ to itself, which sends $1 \in k = T^0E$ to itself. By construction, the composition of $\Phi \circ \Psi$ with the natural mapping from $X$ into $\text{Assoc}_X$ is that mapping. More precisely, the composition of $\Phi \circ \Psi$ with the natural mapping from $X$ into $\text{Assoc}_X$ is the same as the composition of (26.5.3) with the natural mapping from $X$ into $L_X$. This is the same as the natural mapping from $X$ into $\text{Assoc}_X$, as in (26.5.2). It follows that $\Phi \circ \Psi$ is the identity mapping on $\text{Assoc}_X$.

We also have that $\Psi \circ \Phi$ is an algebra homomorphism from $UL_X$ to itself that sends $e_{UL_X}$ to itself. In order to show that $\Psi \circ \Phi$ is the identity mapping on $UL_X$, it suffices to verify that

\[(26.5.6) \quad (\Psi \circ \Phi) \circ i_{UL_X} = i_{UL_X}.\]

Observe that

\[(26.5.7) \quad (\Psi \circ \Phi) \circ i_{UL_X} = \Psi \circ (\Phi \circ i_{UL_X}) = \Psi \circ \phi,\]

by (26.5.3). The composition of this mapping with the natural mapping from $X$ into $L_X$ is the same as the composition of $\Psi$ with the natural mapping from $X$ into $\text{Assoc}_X$, by (26.5.2). This is the same as (26.5.4), as in (26.5.5).

This means that the compositions of both sides of (26.5.6) with the natural mapping from $X$ into $L_X$ are the same. Of course, both sides of (26.5.6) are Lie algebra homomorphisms from $L_X$ into $UL_X$. This implies (26.5.6), because $L_X$ is generated as a Lie algebra over $k$ by the image of the natural mapping from $X$ into $L_X$.

Thus $\Phi$ is an algebra isomorphism from $UL_X$ onto $\text{Assoc}_X$, and $\Psi$ is its inverse. This is the first part of Theorem 4.2 on p20 of [24], which corresponds to Exercise 4 on p95 of [13].

Note that

\[(26.5.8) \quad \phi \text{ maps } L_X \text{ onto the Lie subalgebra of } \text{Assoc}_X \text{ generated by }\]

the image of the natural mapping from $X$ into $\text{Assoc}_X$,

because $\phi$ is a Lie algebra homomorphism from $L_X$ into $\text{Assoc}_X$, and $L_X$ is generated as a Lie algebra over $k$ by the image of the natural mapping from $X$ into $L_X$. This is the easy part of the second part of Theorem 4.2 on p20 of [24].

### 26.6 Injectivity and free modules

Let us continue with the same notation and hypotheses as in the previous section. Under these conditions, we have that

\[(26.6.1) \quad \phi \text{ is injective as a mapping from } L_X \text{ into } \text{Assoc}_X.\]

This means that $\phi$ is a Lie algebra isomorphism from $L_X$ onto the Lie subalgebra of $\text{Assoc}_X$ generated by the image of the natural mapping from $X$ into $\text{Assoc}_X$,
because of (26.5.8). This is the second part of Theorem 4.2 on p20 of [24]. The Lie subalgebra of $\text{Assoc}_X$ generated by the image of the natural mapping from $X$ into $\text{Assoc}_X$ is used to construct free Lie algebras on p95 of [13].

Part of the third part of Theorem 4.2 on p20 of [24] states that

\begin{equation}
L_X \text{ is free as a module over } k.
\end{equation}

Of course, this holds automatically when $k$ is a field. Note that $i_{UL_X}$ is injective as a mapping from $L_X$ into $UL_X$ when (26.6.2) holds, by the Poincaré–Birkhoff–Witt theorem. In this case, (26.6.1) follows from (26.5.3) and the fact that $\Phi$ is an isomorphism from $UL_X$ onto $\text{Assoc}_X$, as mentioned on p21 of [24]. Exercise 6 on p95 of [13] asks how the Poincaré–Birkhoff–Witt theorem is used there, perhaps more precisely in the verification of the requirements of free Lie algebras.

Remember that $L_X$ is graded as an algebra over $k$, as in Section 26.3. If $n \in \mathbb{Z}_+$, then let $L^n_X = L^n_X(k)$ be the submodule of $L_X$, as a module over $k$, consisting of elements of $L_X$ that are homogeneous of degree $n$. The other part of the third part of Theorem 4.2 on p20 of [24] states that

\begin{equation}
L^n_X \text{ is free as a module over } k \text{ for each } n \geq 1.
\end{equation}

This implies (26.6.2), because $L_X$ corresponds to the direct sum of $L^n_X$, $n \geq 1$, as a module over $k$. As before, (26.6.3) holds automatically when $k$ is a field.

If $X$ has only finitely many elements, then $L^n_X$ is finitely generated as a module over $k$ for each $n \geq 1$. More precisely, this follows from the analogous statement for the free algebra $A_X$ on $X$ over $k$. The fourth part of Theorem 4.2 concerns the rank of $L^n_X$, as a free module over $k$, in this case.

Of course, $\text{Assoc}_X = TE$ is graded as an associative algebra over $k$, where $T^nE$ corresponds to the submodule of $TE$ of elements that are homogeneous of degree $n \geq 0$. Let $\text{Assoc}^n_X = \text{Assoc}^n_X(k)$ be the submodule of $\text{Assoc}_X = TE$, as a module over $k$, that corresponds to $T^nE$ for each nonnegative integer $n$. Similarly, $\text{Assoc}_X$ is graded as a Lie algebra over $k$ in this way, with respect to the commutator bracket corresponding to multiplication. Let us check that

\begin{equation}
\phi(L^n_X) \subseteq \text{Assoc}^n_X
\end{equation}

for every $n \geq 1$.

Let $f$ be the natural mapping from $X$ into $\text{Assoc}_X$, which corresponds to the natural inclusion mapping of $X$ into $E = T^1E$. Also let $A_X$ be the free algebra on $X$ over $k$, and let $A^n_X$ be the submodule of $A_X$ consisting of elements that are homogeneous of degree $n \in \mathbb{Z}_+$, as in Section 26.2. There is a unique extension of $f$ to an algebra homomorphism $F$ from $A_X$ into $\text{Assoc}_X$, considered as a Lie algebra over $k$ with respect to the commutator bracket corresponding to multiplication, as before. It is easy to see that

\begin{equation}
F(A^n_X) \subseteq \text{Assoc}^n_X
\end{equation}

for every $n \geq 1$. Remember that $L_X$ is a quotient of $A_X$, and that $L^n_X$ is the image of $A^n_X$ under the quotient mapping. By construction, $F$ is the same as the quotient mapping composed with $\phi$. Thus (26.6.4) follows from (26.6.5).
CHAPTER 26. FREE LIE ALGEBRAS

Suppose that $k$ is a field, and that $X$ has only finitely many elements, as in the first step on p21 of [24]. Let \{\gamma_j\}_{j \in I}$ be a homogeneous basis for $L_X$, as a graded vector space over $k$. This means that for each $j \in I$, $\gamma_j \in L_X^{n(j)}$ for some $n(j) \in \mathbb{Z}_+$, and for each $n \in \mathbb{Z}_+$, the collection of $\gamma_j$ with $n(j) = n$ is a basis for $L_X^n$, as a vector space over $k$. Also let $\preceq$ be a linear ordering on $I$.

The Poincaré–Birkhoff–Witt theorem implies that the family of products of the form
\[
\prod_{i=1}^m \phi(\gamma_{j_i}),
\]
where $j_1, \ldots, j_m \in I$ satisfy $j_1 \preceq \cdots \preceq j_m$ and $m$ is a nonnegative integer, form a basis for $UL_X$. This implies that the corresponding family of products
\[
\phi(\gamma_{j_1}) \cdots \phi(\gamma_{j_m})
\]
is a basis for $\text{Assoc}_X$ as a vector space over $k$. More precisely, this uses (26.5.3) and the fact that $\Phi$ is an algebra isomorphism from $UL_X$ onto $\text{Assoc}_X$. We also have that $\phi(\gamma_j) \in \text{Assoc}_X^{n(j)}$ for every $j \in I$, by (26.6.4). This means that (26.6.7) is homogeneous of degree
\[
n(j_1) + \cdots + n(j_m)
\]
in $\text{Assoc}_X$, which is interpreted as being equal to 0 when $m = 0$.

This can be used to express the dimension of $\text{Assoc}_X$ in terms of the dimensions of $L_X^r$ with $r \leq n$, as in [24]. This means that the dimension of $L_X^r$ can be expressed in terms of the dimension of $\text{Assoc}_X^r$ and the dimensions of $L_X^q$ with $r < n$, as in the remark on p20 of [24]. In particular, the dimension of $L_X^n$ does not depend on $k$.

26.7 Finite sets $X$, $k = \mathbb{Z}$

Let $k$ be a commutative ring with a multiplicative identity element, and let $X$ be a nonempty set. Also let $M_X$ be the free magma on $X$, and let $A_X(k)$ be the free algebra on $X$ over $k$, as in Sections 26.1 and 26.2. Remember that $I_X = I_X(k)$ is the two sided ideal in $A_X(k)$ generated by elements of the form $a a$ and $J(a, b, c)$, where $a, b, c \in A_X(k)$ and $J(a, b, c)$ is as in (26.3.2). Equivalently, $I_X(k)$ is generated as a two-sided ideal in $A_X(k)$ by elements of the form
\[
m n,
\]
with $m \in M_X$,
\[
m_1 m_2 + m_2 m_1 = (m_1 + m_2)(m_1 + m_2) - m_1 m_1 - m_2 m_2,
\]
with $m_1, m_2 \in M_X$, and
\[
J(m_1, m_2, m_3),
\]
with $m_1, m_2, m_3 \in M_X$. This can be seen by expressing $a, b, c \in A_X(k)$ as linear combinations of elements of $M_X$ with coefficients in $k$. 
Let us now consider the cases where $k$ is $\mathbb{Z}$ or $\mathbb{Z}/p\mathbb{Z}$ for some prime number $p$. There is a natural mapping from $A_X(\mathbb{Z})$ onto $A_X(\mathbb{Z}/p\mathbb{Z})$, defined by applying the quotient mapping from $\mathbb{Z}$ onto $\mathbb{Z}/p\mathbb{Z}$ to the coefficients of elements of $A_X(\mathbb{Z})$. One can check that this mapping also sends $I_X(\mathbb{Z})$ onto $I_X(\mathbb{Z}/p\mathbb{Z})$. This leads to a mapping from $L_X(\mathbb{Z})$ onto $L_X(\mathbb{Z}/p\mathbb{Z})$.

More precisely, the mapping from $A_X(\mathbb{Z})$ onto $A_X(\mathbb{Z}/p\mathbb{Z})$ mentioned in the preceding paragraph sends $A^n_X(\mathbb{Z})$ onto $A^n_X(\mathbb{Z}/p\mathbb{Z})$ for every $n \in \mathbb{Z}_+$, where $A^n_X(k)$ consists of the elements of $A_X(k)$ that are homogeneous of degree $n$, as in Section 26.2. Similarly, the induced mapping from $L_X(\mathbb{Z})$ onto $L_X(\mathbb{Z}/p\mathbb{Z})$ sends $L^n_X(\mathbb{Z})$ onto $L^n_X(\mathbb{Z}/p\mathbb{Z})$ for every $n \in \mathbb{Z}_+$, where $L^n_X(k)$ consists of elements of $L_X(k)$ that are homogeneous of degree $n$, as in Section 26.3.

Of course, $p$ times any element of $L_X(\mathbb{Z})$ is mapped to 0 in $L_X(\mathbb{Z}/p\mathbb{Z})$. Conversely, one can check that any element of $L_X(\mathbb{Z})$ that is mapped to 0 in $L_X(\mathbb{Z}/p\mathbb{Z})$ can be expressed as $p$ times an element of $L_X(\mathbb{Z})$.

Suppose now that $X$ has only finitely many elements. Remember that $L^n_X(\mathbb{Z})$ is finitely generated as a module over $\mathbb{Z}$ for every $n \geq 1$ in this case, as in the previous section, which means that $L^n_X(\mathbb{Z})$ is finitely generated as a commutative group. It follows that $L^n_X(\mathbb{Z})$ is isomorphic to a direct sum of finitely many cyclic groups for each $n \geq 1$.

The dimension of $L^n_X(\mathbb{Z}/p\mathbb{Z})$, as a vector space over $\mathbb{Z}/p\mathbb{Z}$, can be determined as in the previous section, and does not depend on $p$. Using this, one can check that $L^n_X(\mathbb{Z})$ is isomorphic to the direct sum of finitely many infinite cyclic groups for each $n \geq 1$. More precisely, the number of factors is the same as the common value of the dimension of $L^n_X(\mathbb{Z}/p\mathbb{Z})$ as a vector space over $\mathbb{Z}/p\mathbb{Z}$.

In particular, $L_X(\mathbb{Z})$ is free as a module over $\mathbb{Z}$ for every $n \geq 1$, which implies that $L_X(\mathbb{Z})$ is free as a module over $\mathbb{Z}$ too. It follows that the mapping $\phi$ from $L_X(\mathbb{Z})$ into $\text{Assoc}_X(\mathbb{Z})$ defined in Section 26.5 is injective, as in the previous section. This corresponds to the second step on p22 of [24].

### 26.8 Arbitrary sets $X$, $k = \mathbb{Z}$

Let $k$ be a commutative ring with a multiplicative identity element, and let $X$ be a nonempty set. Also let $L_X$ be the free Lie algebra on $X$ over $k$, let $E_X$ be the free module on $X$ over $k$, and let $\text{Assoc}_X = T^1E_X$ be the free associative algebra on $X$ over $k$, as in Sections 26.3 and 26.5. Remember that the natural inclusion of $X$ in $E_X = T^1E_X$ leads to a unique Lie algebra homomorphism $\phi = \phi_X$ from $L_X$ into $\text{Assoc}_X$, as before.

Let $Y$ be a nonempty subset of $X$. The free Lie algebra $L_Y$ on $Y$ over $k$, the free module $E_Y$ on $Y$ over $k$, and the free associative algebra $\text{Assoc}_Y = T^1E_Y$ on $Y$ over $k$ can be defined in the same way as before. Of course, there is a natural inclusion of $E_Y$ into $E_X$, which leads to a natural inclusion of $\text{Assoc}_Y$ into $\text{Assoc}_X$. There is a natural Lie algebra homomorphism from $L_Y$ into $L_X$ as well, as in Section 26.3.

Let $\phi_Y$ be the unique Lie algebra homomorphism from $L_Y$ into $\text{Assoc}_Y$ obtained from the natural inclusion of $Y$ in $E_Y$, as in Section 26.5. One can
check that the composition of the natural Lie algebra homomorphism from $L_Y$ into $L_X$ with $\phi_X$ is the same as the composition of $\phi_Y$ with the natural inclusion from $\text{Assoc}_Y$ into $\text{Assoc}_X$.

Suppose now that $k = \mathbb{Z}$. If $Y$ has only finitely many elements, then $\phi_Y$ is injective, as in the previous section. This implies that $\phi_X$ is injective on the image of $L_Y$ in $L_X$ under the natural Lie algebra homomorphism.

It is easy to see that every element of $L_X$ is in the image of $L_Y$ for some nonempty finite subset $Y$ of $X$. It follows that $\phi_X$ is injective as a mapping from $L_X$ into $\text{Assoc}_X$.

Note that $\text{Assoc}_X$ is free as a module over $\mathbb{Z}$, which is to say as a commutative group. It is well known that subgroups of free abelian groups are free abelian groups too.

Using the injectivity of $\phi_X$, we get that $L^n_X(k)$ is a free abelian group for every $n \geq 1$, and that $L_X$ is a free abelian group. Equivalently, these are free as modules over $\mathbb{Z}$. This corresponds to the third step on p22 of [24].

26.9 Arbitrary $X$, $k$

Let $k$ be a commutative ring with a multiplicative identity element, and let $X$ be a nonempty set. Also let $M_X$ be the free magma on $X$, let $A_X(k)$ be the free algebra on $X$ over $k$, as in Sections 26.1, 26.2, and 26.3. Remember that $L_X(k)$ is the quotient of $A_X(k)$ by the two-sided ideal $I_X(k)$ discussed in Sections 26.3 and 26.7. Of course, $A_X(\mathbb{Z})$, $L_X(\mathbb{Z})$, and $I_X(\mathbb{Z})$ can be defined in the same way. If $n \in \mathbb{Z}_+$, then let $L^n_X(k)$, $L^n_X(\mathbb{Z})$ be the subsets of $L_X(k)$, $L_X(\mathbb{Z})$ consisting of elements that are homogeneous of degree $n$, as before.

One can identify $L_X(k)$ with the tensor product of $L_X(\mathbb{Z})$ with $k$ over $\mathbb{Z}$. This is stated more broadly on p19 of [24]. Of course, $A_X(k)$ corresponds to the tensor product of $A_X(\mathbb{Z})$ with $k$ over $\mathbb{Z}$, because $A_X(\mathbb{Z})$ is a free module over $\mathbb{Z}$ with basis $M_X$. We also have that $I_X(k)$ is the image of the tensor product of $I_X(\mathbb{Z})$ with $k$ over $\mathbb{Z}$. Similarly, $L^n_X(k)$ can be identified with the tensor product of $L^n_X(\mathbb{Z})$ with $k$ over $\mathbb{Z}$ for each $n \geq 1$.

Remember that $L_X(\mathbb{Z})$ is a free module over $\mathbb{Z}$ for every $n \geq 1$, and that $L_X(\mathbb{Z})$ is a free module over $\mathbb{Z}$, as in the previous section. It follows that $L^n(k)$ is a free module over $k$ for every $n \geq 1$, and that $L_X(k)$ is a free module over $k$. This implies that the Lie algebra homomorphism $\phi$ from $L_X(k)$ into $\text{Assoc}_X(k)$ defined in Section 26.5 is injective, as in Section 26.6. If $X$ has only finitely many elements, then the rank of $L^n_X(k)$ as a free module over $k$ is the same as the rank of $L^n_X(\mathbb{Z})$ as a free module over $\mathbb{Z}$, which is the same as the dimension in the case of fields. This corresponds to the fourth step on p22 of [24].
26.10   Bases from P. Hall families

Let $X$ be a nonempty set, and let $M_X$ be the free magma on $X$, as in Section 26.1. A P. Hall family in $M_X$ is a subset $H$ of $M_X$ equipped with a linear ordering $\preceq$ that satisfies the following three properties. First,

(26.10.1) $X \subseteq H$.

Second, if $u, v \in H$, then

(26.10.2) $u \preceq v$ when the length of $u$ is strictly less than the length of $v$.

Third, suppose that $u \in M_X \setminus X$, so that $u$ can be expressed in a unique way as

(26.10.3) $u = vw$,

with $v, w \in M_X$. In this case, $u \in H$ if and only if $v, w$ satisfy the following two conditions. The first condition is that

(26.10.4) $v, w \in H$, $v \preceq w$, $v \neq w$.

The second condition is that either $w \in X$, or

(26.10.5) $w = w' w''$, with $w', w'' \in H$, $w' \preceq v$.

This is Definition 5.1 on p22f of [24]. Of course, in (26.10.5), $w$ should satisfy the requirements of the third property of a P. Hall family as well.

Lemma 5.2 on p23 of [24] states that there is always a P. Hall family in $M_X$. Remember that for each positive integer $n$, $X_n$ is the set of elements of $M_X$ of length $n$, as in Section 26.1. We would like to define $H^n \subseteq X_n$ recursively, which will be the set of elements in the Hall family that are in $X_n$. More precisely, we should also choose a linear ordering on $H^n$ for each $n$, which leads to a linear ordering on their union that satisfies (26.10.2). Put $H^1 = X = X_1$, which we can take to be equipped with any linear ordering.

Suppose that $H^1, \ldots, H^{n-1}$ have been chosen for some $n \geq 2$, as well as linear orderings on them. This leads to a linear ordering on their union that satisfies (26.10.2), as before. Suppose also that their union satisfies the three properties of a P. Hall family for elements of length less than or equal to $n - 1$. Using the third property of a P. Hall family, $H^n$ is uniquely determined as a subset of $X_n$. In order to continue the process, we can choose any linear ordering on $H^n$. Thus we can repeat the process, to choose $H^n$ and a linear ordering on it for every $n \geq 1$. It is easy to see that $H = \bigcup_{n=1}^{\infty} H^n$ is a P. Hall family in $M_X$, as desired.

Let $k$ be a commutative ring with a multiplicative identity element, so that the free algebra $A_X(k)$ on $X$ over $k$ can be defined as in Section 26.2. Let $L_X(k)$ be the corresponding free Lie algebra on $X$ over $k$, as in Section 26.3. If $H$ is a P. Hall family in $M_X$, then the elements of $H$ may be considered as elements of $A_X(k)$, which are mapped into $L_X(k)$ by the quotient mapping from $A_X(k)$ onto $L_X(k)$. Theorem 5.3 on p23 of [24] states that the images of the elements of $H$ in $L_X(k)$ form a basis for $L_X(k)$, as a module over $k$. 

26.11 More on \( L^n_X \)

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( X \) be a nonempty set. Thus the free algebra \( A_X \) on \( X \) over \( k \) can be defined as in Section 26.2. Remember that \( A_X \) is graded as an algebra over \( k \) in the strict sense, and let \( A^n_X \) be the submodule of \( A_X \), as a module over \( k \), of elements that are homogeneous of degree \( n \in \mathbb{Z}_+ \), as before. If \( n_1, n_2 \in \mathbb{Z}_+ \), then

\[
(26.11.1) \quad A^n_{X_1} \cdot A^n_{X_2} \subseteq A^{n_1 + n_2}_{X_1},
\]

where the left side is defined as in Section 9.2. If \( n \geq 2 \), then it is easy to see that every element of \( A^n_X \) can be expressed as a finite sum of elements of \( A^n_{X_1} \cdot A^n_{X_2} \), where \( n_1 + n_2 = n \).

Similarly, the free Lie algebra \( L_X \) on \( X \) over \( k \) is a graded algebra over \( k \), as in Section 26.3. Let \( L^n_X \) be the submodule of \( L_X \), as a module over \( k \), consisting of elements that are homogeneous of degree \( n \in \mathbb{Z}_+ \), as before. If \( n_1, n_2 \in \mathbb{Z}_+ \), then

\[
(26.11.2) \quad [L^n_{X_1}, L^n_{X_2}] \subseteq L^{n_1 + n_2}_{X_1},
\]

where the left side is as defined in Section 9.2. If \( n \geq 2 \), then every element of \( L^n_X \) can be expressed as a finite sum of elements of \( [L^n_{X_1}, L^n_{X_2}] \), where \( n_1 + n_2 = n \). This follows from the analogous statement for \( A^n_X \), in the preceding paragraph.

If \( n \geq 2 \), then

\[
(26.11.3) \quad [L^1_X, L^{n-1}_X] = L^n_X,
\]

as in Exercise 2 on p29 of [24]. To see this, let \( n_1, n_2 \) be positive integers with \( n_1 + n_2 = n \), and let \( u_1 \in L^{n_1}_X \) and \( u_2 \in L^{n_2}_X \) be given. We would like to show that \([u_1, u_2] \in [L^1_X, L^{n-1}_X]\), which is obvious when \( n_1 \) or \( n_2 \) is equal to 1. Otherwise, we may as well suppose that \( n_2 \geq n_1 > 1 \). Thus \( u_1 \) can be expressed as a finite sum of terms of the form \([u_1, u_1, u_2, u_2]\), where \( u_{1,1} \in L^{n_1-1}_X \) and \( u_{1,2} \in L^{n_1+2}_X \) for some positive integers \( n_{1,1} \) and \( n_{1,2} \) with \( n_{1,1} + n_{1,2} = n_1 \). This means that \([u_1, u_2] \) can be expressed as a finite sum of terms of the form

\[
(26.11.4) \quad [u_1, u_1, u_2, u_2] = [u_1, u_1, u_1, u_2, u_2] \quad \text{and} \quad [u_1, u_1, u_1, u_2, u_2].
\]

The terms on the right side are elements of \([L^{n_{1,1}+n_{1,2}}_X, L^{n_{1,2}}_X]\) and \([L^{n_{1,1}+n_{1,2}}_X, L^{n_{1,2}}_X]\), respectively. Of course, \( n_{1,1} + n_{1,2} < n_1 \), and so one can repeat the process as needed to get that \([u_1, u_2] \in [L^1_X, L^{n-1}_X]\), as desired. Equivalently, if \( n \geq 2 \), then \( L^n_X \) is generated, as a module over \( k \), by elements of the form

\[
(26.11.5) \quad [x_1, [x_2, \ldots [x_{n-1}, x_n] \ldots]],
\]

where \( x_1, \ldots, x_n \in X \).

Let \( L \) be any Lie algebra over \( k \), and suppose that \( L \) is generated, as a Lie algebra over \( k \), by a nonempty subset \( X \). Under these conditions, \( L \) is generated as a module over \( k \) by the elements of \( X \) and elements of the form

\[(26.11.5) \quad [x_1, [x_2, \ldots [x_{n-1}, x_n] \ldots]],\]

where \( x_1, \ldots, x_n \in X \) for some \( n \geq 2 \). This could be obtained from the previous remarks for free Lie algebras, or verified directly using the same type of arguments.
Chapter 27

Some more Lie algebras

27.1 Generators and relations

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( X \) be a nonempty set. Thus the free Lie algebra \( L_X = L_X(k) \) on \( X \) over \( k \) may be defined as in Section 26.3. If \( E \subseteq L_X \), then let \( R(E) \) be the ideal in \( L_X \) generated by \( E \). Equivalently, \( R(E) \) consists of finite linear combinations of the elements of \( E \) and other elements of \( L_X \) obtained by taking brackets of elements of \( E \) with elements of \( L_X \) any number of times. The quotient \( L_X/R(E) \) may be described as the Lie algebra over \( k \) generated by the images of the elements of \( X \) in the quotient, and with relations given by condition that the images of the elements of \( E \) in the quotient be equal to 0, as on p95 of [13].

If \( B \) is any Lie algebra over \( k \), then any mapping from \( X \) into \( B \) leads to a unique Lie algebra homomorphism from \( L_X \) into \( B \), as before. If this Lie algebra homomorphism maps the elements of \( E \) to 0, then it maps the elements of \( R(E) \) to 0 as well. This leads to an induced Lie algebra homomorphism from \( L_X/R(E) \) into \( B \). If \( B \) is generated as a Lie algebra over \( k \) by the images of the elements of \( X \), then the Lie algebra homomorphism from \( L_X \) into \( B \) just mentioned is surjective. Of course, this implies that the corresponding Lie algebra homomorphism from \( L_X/R(E) \) into \( B \) is surjective as well.

Let \( V \) be a vector space over the real numbers of positive finite dimension, and let \( \Phi \) be a reduced root system in \( V \). If \( \alpha \in \Phi \), then let \( \sigma_\alpha \) be the symmetry on \( V \) with vector \( \alpha \) that maps \( \Phi \) onto itself, as usual. Remember that \( \sigma_\alpha \) can be expressed as \( \sigma_\alpha(v) = v - \lambda_\alpha(v) \alpha \), where \( \lambda_\alpha \) is a linear functional on \( V \) such that \( \lambda_\alpha(\alpha) = 2 \). Let \( (\cdot,\cdot) \) be an inner product on \( V \) that is invariant under the Weyl group of \( \Phi \). If \( \alpha \in \Phi \), then \( \sigma_\alpha \) is the reflection on \( V \) with respect to \( (\cdot,\cdot) \) associated to \( \alpha \), so that \( \lambda_\alpha(\alpha) = 2 (v,\alpha) (\alpha,\alpha)^{-1} \) for every \( v \in V \). Put

\[
(27.1.1) \quad n(\alpha,\beta) = \lambda_\beta(\alpha) = 2 (\alpha,\beta) (\beta,\beta)^{-1}
\]

for every \( \alpha, \beta \in \Phi \), as before. Remember that this is an integer, by the definition of a root system. Let \( \Delta \) be a base for \( \Phi \), so that the restriction of (27.1.1) to
\( \alpha, \beta \in \Delta \) is the Cartan matrix of \( \Phi \) with respect to \( \Delta \). If \( \alpha, \beta \in \Delta \) and \( \alpha \neq \beta \), then
\[
-3 \leq n(\alpha, \beta) \leq 0,
\]
as in Section 20.2. We have also seen that the Cartan matrix is invertible, as in Section 21.15.

Let \( k \) be a field of characteristic 0, and let \( X_\Delta, Y_\Delta, \) and \( H_\Delta \) be disjoint sets, with distinct elements \( x_\alpha, y_\alpha, \) and \( h_\alpha \) indexed by \( \alpha \in \Delta \), respectively. Also let \( \hat{L} = \hat{L}(k) \) be the free Lie algebra on
\[
X_\Delta \cup Y_\Delta \cup H_\Delta
\]
over \( k \). If \( \alpha, \beta \in \Delta \), then put \( \delta_{\alpha, \beta} = 1 \) when \( \alpha = \beta \), and equal to 0 when \( \alpha \neq \beta \).

Let \( \hat{E}_0 \) be the subset of \( \hat{L} \) consisting of the following elements, for \( \alpha, \beta \in \Delta \):
\[
\begin{align*}
\{ [\hat{h}_\alpha, \hat{h}_\beta], \\
[\hat{x}_\alpha, \hat{y}_\beta] - \delta_{\alpha, \beta} \hat{h}_\alpha, \\
[\hat{h}_\alpha, \hat{x}_\beta] - n(\beta, \alpha) \hat{x}_\beta, \\
[\hat{h}_\alpha, \hat{y}_\beta] + n(\beta, \alpha) \hat{y}_\beta.
\end{align*}
\]
This leads to an ideal \( R(\hat{E}_0) \) in \( \hat{L} \), as before.

Let
\[
\tilde{L} = \tilde{L}(k) = \hat{L} / R(\hat{E}_0)
\]
be the corresponding quotient Lie algebra, and let \( \tilde{x}_\alpha, \tilde{y}_\alpha, \) and \( \tilde{h}_\alpha \) be the images of \( x_\alpha, y_\alpha, \) and \( h_\alpha \), respectively, under the quotient mapping for each \( \alpha \in \Delta \). Note that \( \tilde{L} \) is generated by the \( \tilde{x}_\alpha \)’s, \( \tilde{y}_\alpha \)’s, and \( \tilde{h}_\alpha \)’s, \( \alpha \in \Delta \), as a Lie algebra over \( k \). If \( \alpha, \beta \in \Delta \), then we have that
\[
\begin{align*}
[\tilde{h}_\alpha, \tilde{h}_\beta] &= 0, \\
[\tilde{x}_\alpha, \tilde{y}_\beta] &= \delta_{\alpha, \beta} \tilde{h}_\alpha, \\
[\tilde{h}_\alpha, \tilde{x}_\beta] &= n(\beta, \alpha) \tilde{x}_\beta, \\
[\tilde{h}_\alpha, \tilde{y}_\beta] &= -n(\beta, \alpha) \tilde{y}_\beta
\end{align*}
\]
in \( \tilde{L} \), by construction. This Lie algebra has been analyzed by Chevalley, Harish-Chandra, and Jacobson. Here we are largely following the discussion that begins on p96 of [13].

27.2 A module over \( \hat{L} \)

Let us continue with the same notation and hypotheses as in the previous section, and let \( Z_\Delta \) be a vector space over \( k \) with a basis given by distinct elements \( z_\alpha, \alpha \in \Delta \). If \( m \) is a positive integer, then the \( m \)th tensor power \( T^m Z_\Delta \) is a vector space over \( k \), and we can get a basis for \( T^m Z_\Delta \) using elements of the form \( z_{\alpha_1} \otimes \cdots \otimes z_{\alpha_m} \), where \( \alpha_1, \ldots, \alpha_m \in \Delta \). Remember that the tensor algebra \( TZ_\Delta \)
27.2. A MODULE OVER $\hat{L}$

is obtained by taking the direct sum of $T^m \mathbb{Z}_\Delta$ over all nonnegative integers $m$, with $T^0 \mathbb{Z}_\Delta = k$, as in Section 25.3. We would like to make $T \mathbb{Z}_\Delta$ into a module over $\hat{L}$ in a certain way, as on p97 of [13]. To do this, we need to define the action of the generators of $\hat{L}$ on $T \mathbb{Z}_\Delta$.

Of course, we can get a basis for $T \mathbb{Z}_\Delta$, as a vector space over $k$, using the bases for $T^m \mathbb{Z}_\Delta$ mentioned in the preceding paragraph for each $m \geq 1$, and $1 \in k$. In order to define the action of the generators of $\hat{L}$ on $T \mathbb{Z}_\Delta$, it suffices to define the action on the basis elements of $T \mathbb{Z}_\Delta$. If $\alpha \in \Delta$, then we put

$$(27.2.1) \quad \hat{h}_\alpha \cdot 1 = 0,$$

and

$$(27.2.2) \quad \hat{h}_\alpha \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m}) = -(n(\gamma_1, \alpha) + \cdots + n(\gamma_m, \alpha)) (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m})$$

for every $m \geq 1$ and $\gamma_1, \ldots, \gamma_m \in \Delta$. Similarly, we put

$$(27.2.3) \quad \hat{y}_\alpha \cdot 1 = z_\alpha,$$

and

$$(27.2.4) \quad \hat{y}_\alpha \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m}) = z_\alpha \otimes z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m}$$

for every $m \geq 1$ and $\gamma_1, \ldots, \gamma_m \in \Delta$. We also put

$$(27.2.5) \quad \hat{x}_\alpha \cdot 1 = 0,$$

and

$$(27.2.6) \quad \hat{x}_\alpha \cdot z_\gamma = 0$$

for every $\gamma \in \Delta$. If $m \geq 2$ and $\gamma_1, \ldots, \gamma_m \in \Delta$, then we put

$$(27.2.7) \quad \hat{x}_\alpha \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m}) = z_{\gamma_1} \otimes (\hat{x}_\alpha \cdot (z_{\gamma_2} \otimes \cdots \otimes z_{\gamma_m})) - \delta_{\gamma_1, \alpha} (n(\gamma_2, \alpha) + \cdots + n(\gamma_m, \alpha)) (z_{\gamma_2} \otimes \cdots \otimes z_{\gamma_m}).$$

More precisely, this can be used to define the left side recursively.

Let $\text{gl}(T \mathbb{Z}_\Delta)$ be the space of linear mappings from $T \mathbb{Z}_\Delta$ into itself, as a Lie algebra over $k$ with respect to the corresponding commutator bracket, as usual. If $\alpha \in \Delta$, then we get elements of $\text{gl}(T \mathbb{Z}_\Delta)$ associated to $\hat{h}_\alpha$, $\hat{y}_\alpha$, and $\hat{x}_\alpha$, as in the previous paragraph. This leads to a Lie algebra homomorphism $\hat{\phi}$ from $\hat{L}$ into $\text{gl}(T \mathbb{Z}_\Delta)$, so that $T \mathbb{Z}_\Delta$ becomes a module over $\hat{L}$, as a Lie algebra over $k$. We would like to show that

$$(27.2.8) \quad \mathcal{R}(\hat{E}_0) \subseteq \ker \hat{\phi},$$

where $\mathcal{R}(\hat{E}_0)$ is as in the previous section. This corresponds to the proposition on p97 of [13].
Of course, it suffices to show that
\[ \tilde{E}_0 \subseteq \ker \tilde{\phi}, \]
where \( \tilde{E}_0 \) is as in the previous section, because \( \mathcal{R}(\tilde{E}_0) \) is the ideal in \( \hat{L} \) generated by \( \tilde{E}_0 \). If \( \alpha, \beta \in \Delta \), then it is easy to see that
\[ \phi([\hat{h}_\alpha, \hat{h}_\beta]) = [\tilde{\phi}(\hat{h}_\alpha), \tilde{\phi}(\hat{h}_\beta)] = 0, \]
which is to say that the actions of \( \hat{h}_\alpha \) and \( \hat{h}_\beta \) on \( T\mathcal{Z}_\Delta \) commute. Note that the action of \( \hat{y}_\beta \) on \( T\mathcal{Z}_\Delta \) is the same as multiplication by \( z_\alpha \) on the left, with respect to multiplication on the tensor algebra \( T\mathcal{Z}_\Delta \).

Let \( \alpha, \beta \in \Delta \) be given, and let us check that
\[ \phi([\hat{h}_\alpha, \hat{y}_\beta]) = [\tilde{\phi}(\hat{h}_\alpha), \tilde{\phi}(\hat{y}_\beta)] = 0, \]
which is the same as \( \delta_{\alpha, \beta} \hat{h}_\alpha \cdot 1 \). If \( m \geq 2 \) and \( \gamma_1, \ldots, \gamma_m \in \Delta \), then
\[ \hat{x}_\alpha \cdot (\hat{y}_\beta \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m})) - \hat{y}_\beta \cdot (\hat{x}_\alpha \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m})) \]
\[ = -\delta_{\alpha, \beta} \cdot (\alpha(\gamma_2, \alpha) + \cdots + \alpha(\gamma_m, \alpha)) \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m}) \]
\[ = \delta_{\beta, \alpha} \hat{h}_\alpha \cdot (z_{\gamma_2} \otimes \cdots \otimes z_{\gamma_m}). \]
More precisely, this uses (27.2.7) in the second step. This shows that (27.2.11) holds.

Similarly, let us verify that
\[ \phi([\hat{h}_\alpha, \hat{y}_\beta]) = [\tilde{\phi}(\hat{h}_\alpha), \tilde{\phi}(\hat{y}_\beta)] = -n(\beta, \alpha) \tilde{\phi}(\hat{y}_\beta). \]
Observe that
\[ \hat{h}_\alpha \cdot (\hat{y}_\beta \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m})) - \hat{y}_\beta \cdot (\hat{h}_\alpha \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m})) \]
\[ = -n(\beta, \alpha) z_\beta = -n(\beta, \alpha) \hat{y}_\beta \cdot 1. \]
If \( \gamma_1, \ldots, \gamma_m \in \Delta \) for some \( m \geq 1 \), then
\[ \hat{h}_\alpha \cdot (\hat{y}_\beta \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m})) - \hat{y}_\beta \cdot (\hat{h}_\alpha \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m})) \]
\[ = -n(\beta, \alpha) (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m}) \]
\[ = -n(\beta, \alpha) \hat{y}_\beta \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m}). \]
This implies (27.2.14).
Note that
\(\hat{h}_\alpha \cdot (\hat{x}_\beta \cdot 1) = n(\beta, \alpha) \hat{x}_\beta \cdot 1 = 0.\)

We would like to check that
\[
\hat{h}_\alpha \cdot (\hat{x}_\beta \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m}))
= - (n(\gamma_1, \alpha) + \cdots + n(\gamma_m, \alpha) - n(\beta, \alpha)) \hat{x}_\beta \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m})
\]
for every \(m \geq 1\) and \(\gamma_1, \ldots, \gamma_m \in \Delta\). If \(m = 1\), then both sides of the equation are equal to 0. Suppose that \(m \geq 2\), and that the analogous statement holds when \(m - 1\), by induction, so that
\[
\hat{h}_\alpha \cdot (\hat{x}_\beta \cdot (z_{\gamma_2} \otimes \cdots \otimes z_{\gamma_m}))
= - (n(\gamma_2, \alpha) + \cdots + n(\gamma_m, \alpha) - n(\beta, \alpha)) \hat{x}_\beta \cdot (z_{\gamma_2} \otimes \cdots \otimes z_{\gamma_m}).
\]

Using this, one can verify that
\[
\hat{h}_\alpha \cdot (\hat{x}_\beta \cdot (z_{\gamma_1} \otimes (\hat{x}_\beta \cdot (z_{\gamma_2} \otimes \cdots \otimes z_{\gamma_m}))))
= - (n(\gamma_1, \alpha) + \cdots + n(\gamma_m, \alpha) - n(\beta, \alpha)) (z_{\gamma_1} \otimes (\hat{x}_\beta \cdot (z_{\gamma_2} \otimes \cdots \otimes z_{\gamma_m}))).
\]

One can obtain (27.2.18) from this and (27.2.7). More precisely, this also uses the fact that
\[
\hat{h}_\alpha \cdot (\delta_{\gamma_1, \beta} (z_{\gamma_2} \otimes \cdots \otimes z_{\gamma_m}))
= - \delta_{\gamma_1, \beta} (n(\gamma_2, \alpha) + \cdots + n(\gamma_m, \alpha)) (z_{\gamma_2} \otimes \cdots \otimes z_{\gamma_m})
= - \delta_{\gamma_1, \beta} (n(\gamma_1, \alpha) + n(\gamma_2, \alpha) + \cdots + n(\gamma_m, \alpha) - n(\beta, \alpha)) (z_{\gamma_2} \otimes \cdots \otimes z_{\gamma_m}).
\]

Let us check that
\[
\hat{\phi}(\hat{h}_\alpha, \hat{x}_\beta) = [\hat{\phi}(\hat{h}_\alpha), \hat{\phi}(\hat{x}_\beta)] = n(\beta, \alpha) \hat{\phi}(\hat{x}_\beta).
\]

Observe that
\[
\hat{h}_\alpha \cdot (\hat{x}_\beta \cdot 1) - \hat{x}_\beta \cdot (\hat{h}_\alpha \cdot 1) = 0,
\]
which is the same as \(n(\beta, \alpha) \hat{x}_\beta \cdot 1\). Let \(\gamma_1, \ldots, \gamma_m \in \Delta\) be given for some \(m \geq 1\), and note that
\[
\hat{x}_\beta \cdot (\hat{h}_\alpha \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m}))
= - (n(\gamma_1, \alpha) + \cdots + n(\gamma_m, \alpha)) \hat{x}_\beta \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m}).
\]

Combining this with (27.2.18), we obtain that
\[
\hat{h}_\alpha \cdot (\hat{x}_\beta \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m})) - \hat{x}_\beta \cdot (\hat{h}_\alpha \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m}))
= (- (n(\gamma_1, \alpha) + \cdots + n(\gamma_m, \alpha) - n(\beta, \alpha)) + (n(\gamma_1, \alpha) + \cdots + n(\gamma_m, \alpha)) \hat{x}_\beta \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m})
= n(\beta, \alpha) \hat{x}_\beta \cdot (z_{\gamma_1} \otimes \cdots \otimes z_{\gamma_m}).
\]

This implies (27.2.22), using (27.2.23) as well.

It follows that (27.2.9) holds, because of (27.2.10), (27.2.11), (27.2.14), and (27.2.22). This implies (27.2.8), as before.
27.3 Some properties of $\tilde{L}$

We continue with the same notation and hypotheses as in the previous two sections. In particular, remember that $\tilde{L} = \hat{L}/R(\hat{E}_0)$, and that $\tilde{x}_\alpha, \tilde{y}_\alpha, \tilde{h}_\alpha$ are the images of $x_\alpha, y_\alpha, h_\alpha \in \hat{L}$ in $\hat{L}$, respectively, under the natural quotient mapping for each $\alpha \in \Delta$. In the previous section, a Lie algebra homomorphism $\tilde{\phi}$ from $\hat{L}$ into $gl(TZ_\Delta)$ was defined, whose kernel contains $R(\hat{E}_0)$. This leads to a Lie algebra homomorphism $e\phi$ from $eL$ into $gl(TZ_\Delta)$, whose composition with the natural quotient mapping from $bL$ onto $eL$ is equal to $b\phi$, as in the proposition on p97 of [13]. Thus $TZ_\Delta$ may be considered as a module over $eL$, as a Lie algebra over $k$.

Of course, $e\phi(x) = b\phi(x)$, $e\phi(y) = b\phi(y)$, and $e\phi(h) = b\phi(h)$ for every $\alpha \in \Delta$, by construction. This means that the actions of $x_\alpha, y_\alpha,$ and $h_\alpha$ on $TZ_\Delta$ are the same as for $\tilde{x}_\alpha, \tilde{y}_\alpha,$ and $\tilde{h}_\alpha$, respectively, for each $\alpha \in \Delta$.

Let $eH$ be the linear span of the $e_h$'s, $\alpha \in \Delta$, in $eL$. This is a Lie subalgebra of $L$, that is commutative as a Lie algebra, because the $e_h$'s commute with each other, as in (27.1.9). Let $eX$ be the Lie subalgebra of $eL$ generated by the $e_x$'s, $\alpha \in \Delta$, and let $eY$ be the Lie subalgebra of $eL$ generated by the $e_y$'s, $\alpha \in \Delta$. The theorem on p97f of [13] states that

(27.3.1) the $\tilde{h}_\alpha$'s, $\alpha \in \Delta,$ form a basis for $\tilde{H}$ as a vector space over $k$, and that

(27.3.2) $\tilde{L}$ corresponds to the direct sum of $\tilde{X}$, $\tilde{H}$, and $\tilde{Y}$, as a vector space over $k$.

Step (1) on p98 of [13] states that the intersection of the linear span of the $\tilde{h}_\alpha$'s, $\alpha \in \Delta$, in $\tilde{L}$ with the kernel of $\tilde{\phi}$ is trivial. To see this, let $a_\alpha \in k$ be given for each $\alpha \in \Delta$, and put

(27.3.3) $\tilde{\alpha} = \sum_{\alpha \in \Delta} a_\alpha \tilde{h}_\alpha$.

Thus $\tilde{\alpha} \in \tilde{L}$, and

(27.3.4) $\tilde{\alpha} \cdot z_\gamma = -\sum_{\alpha \in \Delta} a_\alpha n(\gamma, \alpha) z_\gamma$

for every $\gamma \in \Delta$, by (27.2.2). If $\tilde{\phi}(\tilde{\alpha}) = 0$, then (27.3.4) is equal to 0 for each $\gamma \in \Delta$, so that

(27.3.5) $\sum_{\alpha \in \Delta} a_\alpha n(\gamma, \alpha) = 0$

for every $\gamma \in \Delta$. We would like to use this to get that $a_\alpha = 0$ for every $\alpha \in \Delta$.

Remember that the Cartan matrix is invertible, as a matrix of real numbers, as in Section 21.15. The entries of the Cartan matrix are integers, so that its determinant is a nonzero integer. If we consider the Cartan matrix as having entries in $k$, then its determinant is still nonzero, because $k$ has characteristic 0,
27.3. SOME PROPERTIES OF $\tilde{L}$

by hypothesis. This means that the Cartan matrix is also invertible as a matrix with entries in $k$. Thus (27.3.5) implies that $a_\alpha = 0$ for every $\alpha \in \Delta$, as desired.

Of course,

$$\tilde{h} = \sum_{\alpha \in \Delta} a_\alpha \tilde{h}_\alpha$$

(27.3.6)

is the image of (27.3.3) under the natural quotient mapping from $\tilde{L}$ onto $\tilde{L}$. If $\tilde{h} = 0$, then $\phi(\tilde{h}) = \phi(\tilde{h}) = 0$, so that $a_\alpha = 0$ for every $\alpha \in \Delta$, as before. It follows that the restriction of the natural quotient mapping from $\tilde{L}$ onto $\tilde{L}$ to the linear span of the $\tilde{h}_\alpha$'s, $\alpha \in \Delta$, in $\tilde{L}$ is injective, as in Step (2) on p98 of [13]. This implies that the dimension of $\tilde{H}$, as a vector space over $k$, is the same as the number of elements in $\Delta$, as in Step (4) on p98 of [13]. Equivalently, this means that (27.3.1) holds.

In particular,

$$\tilde{h}_\alpha \neq 0 \text{ for each } \alpha \in \Delta.$$  

(27.3.7)

This implies that

$$\tilde{x}_\alpha, \tilde{y}_\alpha \neq 0 \text{ for each } \alpha \in \Delta,$$

(27.3.8)

because $[\tilde{x}_\alpha, \tilde{y}_\alpha] = \tilde{h}_\alpha$, as in (27.1.10).

If $\beta \in \Delta$, then put

$$f_\beta(\alpha) = n(\beta, \alpha)$$

(27.3.9)

for every $\alpha \in \Delta$, which may be considered as a $k$-valued function on $\Delta$. In fact, the $f_\beta$'s form a basis for the space of $k$-valued functions on $\Delta$, as a vector space over $k$. This follows from the invertibility of the Cartan matrix, as a matrix with entries in $k$, as before. In particular, for each $\beta \in \Delta$, $f_\beta \neq 0$ on $\Delta$. Similarly, if $\beta, \beta' \in \Delta$ and $\beta \neq \beta'$, then $f_\beta \neq f_{\beta'}$, as functions on $\Delta$. We also have that $f_\beta \neq -f_{\beta'}$, even when $\beta = \beta'$. This uses the hypothesis that $k$ has characteristic 0, so that $1 \neq -1$ in $k$.

One can use this to check that

the set of $\tilde{x}_\beta$'s, $\tilde{y}_\beta$'s, and $\tilde{h}_\beta$'s, $\beta \in \Delta$,

is linearly independent in $\tilde{L}$,

(27.3.10)

as in Step (3) on p98 of [13]. This also uses the fact that the $\tilde{x}_\beta$'s, $\tilde{y}_\beta$'s, and $\tilde{h}_\beta$'s are simultaneous eigenvectors for $\text{ad}_{\tilde{h}_\alpha}$, $\alpha \in \Delta$, as in (27.1.9), (27.1.11), and (27.1.12), with the corresponding eigenvalues given by $f_\beta(\alpha)$, $-f_{\beta'}(\alpha)$, and 0, respectively. Note that the $\text{ad}_{\tilde{h}_\alpha}$'s, $\alpha \in \Delta$, commute with each other, as linear mappings on $\tilde{L}$, because the $\tilde{h}_\alpha$'s commute in $\tilde{L}$, as in Section 2.4.

More precisely, suppose that we have a linear combination of the $\tilde{x}_\beta$'s, $\tilde{y}_\beta$'s, and $\tilde{h}_\beta$'s, $\beta \in \Delta$, in $\tilde{L}$ that is equal to 0. One can show that the coefficients of the $\tilde{x}_\beta$'s and $\tilde{y}_\beta$'s are all equal to 0, and that the corresponding linear combination of the $\tilde{h}_\beta$'s is equal to 0, using the earlier remarks about the $f_\beta$'s. It follows that the coefficients of the $\tilde{h}_\beta$'s have to be equal to 0 too, by (27.3.1).
CHAPTER 27. SOME MORE LIE ALGEBRAS

Equivalently, consider the linear subspace of $\hat{L}$ spanned by the $\hat{x}_\beta$’s, $\hat{y}_\beta$’s, and $\hat{h}_\gamma$’s, $\beta \in \Delta$. The restriction of the natural quotient mapping from $\hat{L}$ onto $\hat{L}$ to this linear subspace is injective, as in Step (3) on p98 of [13].

27.4 Some brackets in $\hat{L}$

Let us continue with the discussion in the previous section. In particular, we would like to prove (27.3.2).

If $\tilde{u}_1, \ldots, \tilde{u}_r \in \hat{L}$ for some $r \geq 2$, then put

$$[\tilde{u}_1, \ldots, \tilde{u}_r] = [\tilde{u}_1, [\tilde{u}_2, \ldots, [\tilde{u}_{r-1}, \tilde{u}_r] \ldots]].$$

(27.4.1)

If $\alpha, \beta_1, \ldots, \beta_r \in \Delta$, then

$$[\hat{h}_\alpha, [\hat{x}_{\beta_1}, \ldots, \hat{x}_{\beta_r}]] = (n(\beta_1, \alpha) + \cdots + n(\beta_r, \alpha)) [\hat{x}_{\beta_1}, \ldots, \hat{x}_{\beta_r}],$$

(27.4.2)

as in Step (5) on p98 of [13]. This follows from (27.1.11) and the Jacobi identity when $r = 2$, and one can use induction to get the analogous conclusion for every $r \geq 2$. Similarly,

$$[\hat{h}_\alpha, [\hat{y}_{\beta_1}, \ldots, \hat{y}_{\beta_r}]] = -(n(\beta_1, \alpha) + \cdots + n(\beta_r, \alpha)) [\hat{y}_{\beta_1}, \ldots, \hat{y}_{\beta_r}],$$

(27.4.3)

by (27.1.12).

If $\alpha, \beta_1, \ldots, \beta_r \in \Delta$ for some $r \geq 2$ again, then

$$[\hat{y}_\alpha, [\hat{x}_{\beta_1}, \ldots, \hat{x}_{\beta_r}]]$$

(27.4.4)

can be expressed as a linear combination of $\hat{x}_{\beta_1}$ and $\hat{x}_{\beta_2}$ when $r = 2$, and as a linear combination of terms of the form

$$[\hat{x}_{\gamma_1}, \ldots, \hat{x}_{\gamma_{r-1}}]$$

(27.4.5)

when $r \geq 3$, where $\gamma_1, \ldots, \gamma_{r-1} \in \{\beta_1, \ldots, \beta_r\}$. In particular, this means that (27.4.4) is an element of $\hat{X}$, as in Step 6 on p98 of [13]. If $r = 2$, then the previous statement can be obtained from (27.1.10), (27.1.11), and the Jacobi identity. If $r \geq 3$, then one can use induction and (27.4.2). Similarly,

$$[\hat{x}_\alpha, [\hat{y}_{\beta_1}, \ldots, \hat{y}_{\beta_r}]]$$

(27.4.6)

can be expressed as a linear combination of $\hat{y}_{\beta_1}$ and $\hat{y}_{\beta_2}$ when $r = 2$, and as a linear combination of terms of the form

$$[\hat{y}_{\gamma_1}, \ldots, \hat{y}_{\gamma_{r-1}}]$$

(27.4.7)

when $r \geq 3$, where $\gamma_1, \ldots, \gamma_{r-1} \in \{\beta_1, \ldots, \beta_r\}$, as before.

Remember that $\hat{X}$ and $\hat{Y}$ are the Lie subalgebras of $\hat{L}$ generated by the $\hat{x}_\alpha$’s and $\hat{y}_\alpha$’s, respectively, with $\alpha \in \Delta$. Note that $\hat{X}$ is spanned by the $\hat{x}_\alpha$’s, $\alpha \in \Delta$, and elements of the form $[\hat{x}_{\alpha_1}, \ldots, \hat{x}_{\alpha_r}]$, where $\alpha_1, \ldots, \alpha_r \in \Delta$ for some $r \geq 2$,
27.5. Eigenvectors in $\tilde{L}$

Let us look a bit more at the simultaneous eigenvectors for $\text{ad}_{\tilde{h}_\alpha}$, $\alpha \in \Delta$, in $\tilde{L}$, as in the previous two sections. Remember that $\tilde{H}$ is the linear span of the $\tilde{h}_\alpha$'s, $\alpha \in \Delta$. Let $\tilde{H}'$ be the dual of $\tilde{H}$, as a vector space over $k$, as usual. If $\tilde{\mu} \in \tilde{H}'$, then put

$$(27.5.1) \quad \tilde{L}_{\tilde{\mu}} = \{ \tilde{u} \in \tilde{L} : [\tilde{h}, \tilde{u}] = \tilde{\mu}(\tilde{h}) \tilde{u} \text{ for every } \tilde{h} \in \tilde{H} \},$$

which is a linear subspace of $\tilde{L}$. Of course, the $\text{ad}_{\tilde{h}_\alpha}$'s, $\tilde{h} \in \tilde{H}$, commute as linear mappings on $\tilde{L}$, because of the same property of the $\text{ad}_{h_\alpha}$'s, $\alpha \in \Delta$.

Suppose for the moment that $\tilde{\mu}_1, \ldots, \tilde{\mu}_n$ are distinct elements of $\tilde{H}'$, and that $\tilde{u}_j \in \tilde{L}_{\tilde{\mu}_j}$ for $j = 1, \ldots, n$. If $\sum_{j=1}^{n} \tilde{u}_j = 0$, then $\tilde{u}_j = 0$ for every $j = 1, \ldots, n$, by standard arguments.

Remember that the $\tilde{h}_\alpha$'s, $\alpha \in \Delta$, form a basis for $\tilde{H}$. If $\beta \in \Delta$, then there is a unique $\tilde{\nu}_\beta \in \tilde{H}'$ such that

$$(27.5.2) \quad \tilde{\nu}_\beta(\tilde{h}_\alpha) = n(\beta, \alpha)$$

for every $\alpha \in \Delta$. The $\tilde{\nu}_\beta$'s, $\beta \in \Delta$, form a basis for $\tilde{H}'$, as a vector space over $k$. This is essentially the same as the fact that the $f_\beta$'s in (27.3.9) form a basis for the space of $k$-valued functions on $\Delta$, as before.

Every element of $\tilde{X}$ can be expressed as the sum of elements of finitely many of the subspaces $\tilde{L}_{\tilde{\mu}_j}$, where $\tilde{\mu} \in \tilde{H}'$ can be expressed as a nontrivial sum of $\tilde{\nu}_\beta$'s,
as in the previous section. Similarly, every element of \( \tilde{Y} \) can be expressed as the sum of finitely many elements of \( \tilde{L}_\mu' \)'s, where \( \tilde{\mu} \in \tilde{H}' \) can be expressed as a nontrivial sum of \( -\tilde{\nu}_\beta \)'s. It follows that \( \tilde{X} \) and \( \tilde{Y} \) correspond to direct sums of their intersections with the appropriate subspaces \( \tilde{L}_\mu' \), as vector spaces over \( k \).

Clearly \( \tilde{H} \subseteq \tilde{L}_0 \), by (27.1.9). Remember that \( \tilde{L} \) corresponds to the direct sum of \( \tilde{X}, \tilde{H}, \) and \( \tilde{Y} \), as a vector space over \( k \), as in (27.3.2). Using this, one can check that

\[
\tilde{H} = \tilde{L}_0.
\]

This also uses the hypothesis that \( k \) has characteristic 0, so that nontrivial sums of \( \tilde{\nu}_\beta \)'s cannot be equal to 0.

Similarly, \( \tilde{L}_\mu' \subseteq \tilde{X} \) when \( \tilde{\mu} \in \tilde{H} \) is a nontrivial sum of \( \tilde{\nu}_\beta \)'s, and \( \tilde{L}_\mu' \subseteq \tilde{Y} \) when \( \tilde{\mu} \) is a nontrivial sum of \( -\tilde{\nu}_\beta \)'s. This uses the hypothesis that \( k \) has characteristic 0, to get that nontrivial sums of \( \tilde{\nu}_\beta \)'s cannot be equal to nontrivial sums of \( -\tilde{\nu}_\beta \)'s. This implies that \( \tilde{X} \) corresponds to the direct sum of the subspaces \( \tilde{L}_\mu' \), where \( \tilde{\mu} \in \tilde{H}' \) can be expressed as a nontrivial sum of \( \tilde{\nu}_\beta \)'s, and that \( \tilde{Y} \) corresponds to the direct sum of the suspaces \( \tilde{L}_\mu' \), where \( \tilde{\mu} \in \tilde{H}' \) can be expressed as a nontrivial sum of \( -\tilde{\nu}_\beta \)'s, as vector spaces over \( k \).

Let \( \tilde{\mu} \in \tilde{H}' \) be given, and suppose that \( \tilde{L}_\mu' \neq \{0\} \). Under these conditions, one can verify that either \( \tilde{\mu} = 0 \), or \( \tilde{\mu} \) is a nontrivial sum of \( \tilde{\nu}_\beta \)'s, or \( \tilde{\mu} \) is a nontrivial sum of \( -\tilde{\nu}_\beta \)'s.

If \( \tilde{\mu} \in \tilde{H}' \), then \( \tilde{L}_\mu' \) is finite-dimensional, as a vector space over \( k \). This follows from (27.5.3) when \( \tilde{\mu} = 0 \), and otherwise we may suppose that \( \tilde{\mu} \) is a nontrivial sum of \( \nu_\beta \)'s, or that \( \tilde{\mu} \) is a nontrivial sum of \( -\nu_\beta \)'s, as in the preceding paragraph. In both cases, the elements of \( \tilde{L}_\mu' \) considered in the previous section are contained in finite-dimensional subspaces of \( \tilde{L} \). One can check that all of the elements of \( \tilde{L}_\mu' \) can be obtained as linear combinations of those considered before, because \( \tilde{L} \) is spanned by elements of this type and the \( \tilde{h}_\alpha \)'s, as in the previous section.

This corresponds to some remarks on p98 of [13], after the proof of the theorem, and on p53 of [23]. In particular, the elements of \( \tilde{L}_\mu' \) are said to have weight \( \tilde{\mu} \).

### 27.6 Some more properties of \( \tilde{L} \)

Let us continue with the same notation and hypotheses as in the previous five sections. Thus the tensor algebra \( T\Delta \) from Section 27.2 is an associative algebra over \( k \), which may be considered as a Lie algebra over \( k \) with respect to the corresponding commutator bracket. Let \( L\Delta \) be the free Lie algebra on \( Y_\Delta = \{ \tilde{y}_\alpha : \alpha \in \Delta \} \) over \( k \), and let \( \psi_1 \) be the Lie algebra homomorphism from \( L\Delta \) into \( T\Delta \) that sends \( \tilde{y}_\alpha \) to \( z_\alpha \in Z\Delta \) for every \( \alpha \in \Delta \). Note that

\[
\psi_1 \text{ is injective as a mapping from } L\Delta \text{ into } T\Delta,
\]
as in Section 26.6. More precisely, this corresponds to identifying $TZ_\Delta$ with the free associative algebra $\text{Assoc}_{Y_\Delta}$ on $Y_\Delta$ over $k$, by identifying $\tilde{y}_\alpha$ with $z_\alpha$ for each $\alpha \in \Delta$.

Of course, each element of $TZ_\Delta$ determines a linear mapping from $TZ_\Delta$ into itself, defined by multiplication on the left by the given element of $TZ_\Delta$. In particular, this defines a Lie algebra homomorphism from $TZ_\Delta$ into $gl(TZ_\Delta)$. Note that this mapping is injective, because $TZ_\Delta$ has a multiplicative identity element. Let $\psi'_1$ be the mapping from $L_{Y_\Delta}$ into $gl(TZ_\Delta)$ which is the composition of $\psi_1$ with the Lie algebra homomorphism from $TZ_\Delta$ into $gl(TZ_\Delta)$ just mentioned. Thus $\psi'_1$ is an injective Lie algebra homomorphism from $L_{Y_\Delta}$ into $gl(TZ_\Delta)$.

Let $\tilde{\phi}$ be the Lie algebra homomorphism from $\hat{L}$ into $gl(TZ_\Delta)$ defined in Section 27.2. Remember that $\tilde{\phi}(\tilde{y}_\alpha)$ is the linear mapping from $TZ_\Delta$ to itself corresponding to left multiplication by $z_\alpha$ for each $\alpha \in \Delta$. There is a natural Lie algebra homomorphism from $L_{Y_\Delta}$ into $\hat{L}$, which sends $\tilde{y}_\alpha$ to itself as an element of $\hat{L}$ for every $\alpha \in \Delta$. Let $\tilde{\phi}_1$ be the composition of this homomorphism with $\tilde{\phi}$, so that $\tilde{\phi}_1$ is a Lie algebra homomorphism from $L_{Y_\Delta}$ into $gl(TZ_\Delta)$. Observe that

$$\tilde{\phi}_1 = \psi'_1$$

on $L_{Y_\Delta}$, because both mappings send $\tilde{y}_\alpha$ to the linear mapping from $TZ_\Delta$ into itself that corresponds to left multiplication by $z_\alpha$ for each $\alpha \in \Delta$.

Remember from Section 27.3 that $\tilde{\phi}$ is the same as the composition of the natural quotient mapping from $\hat{L}$ onto $\hat{L}$ with the Lie algebra homomorphism $\phi$ from $L$ into $gl(TZ_\Delta)$. Let $\rho_1$ be the composition of the natural mapping from $L_{Y_\Delta}$ into $\hat{L}$ with the quotient mapping from $\hat{L}$ onto $\hat{L}$. Thus $\rho_1$ is a Lie algebra homomorphism from $L_{Y_\Delta}$ into $\hat{L}$, and

$$\tilde{\phi}_1 = \tilde{\phi} \circ \rho_1$$

on $L_{Y_\Delta}$, by construction. It follows that

$$\rho_1$$

is injective as a mapping from $L_{Y_\Delta}$ into $\hat{L}$, because of (27.6.2) and the fact that $\psi'_1$ is injective on $L_{Y_\Delta}$, as before.

Note that $\rho_1(\tilde{y}_\alpha) = \tilde{y}_\alpha$ for each $\alpha \in \Delta$, by construction. Let $\tilde{Y}$ be the Lie subalgebra of $\hat{L}$ generated by the $\tilde{y}_\alpha$’s, $\alpha \in \Delta$, as in Section 27.3. Thus

$$\rho_1(L_{Y_\Delta}) = \tilde{Y},$$

because $L_{Y_\Delta}$ is generated by $Y_\Delta$ as a Lie algebra over $k$. This shows that $\tilde{Y}$ is isomorphic to $L_{Y_\Delta}$ as a Lie algebra over $k$, as in Exercise 1 on p101 of [13].

Let $\tilde{\sigma}$ be the Lie algebra homomorphism from $\hat{L}$ into itself such that

$$\tilde{\sigma}(\tilde{x}_\alpha) = -\tilde{y}_\alpha, \quad \tilde{\sigma}(\tilde{h}_\alpha) = -\tilde{h}_\alpha$$

for every $\alpha \in \Delta$. The composition of $\tilde{\sigma}$ with itself sends $\tilde{x}_\alpha$, $\tilde{y}_\alpha$, and $\tilde{h}_\alpha$ to themselves, respectively, for every $\alpha \in \Delta$. Thus $\tilde{\sigma} \circ \tilde{\sigma}$ is the identity mapping.
CHAPTER 27. SOME MORE LIE ALGEBRAS

on \( \hat{L} \), and in particular \( \hat{\sigma} \) is a Lie algebra automorphism of \( \hat{L} \). It is easy to see that
\[
\hat{\sigma}(\mathcal{R}(\hat{E}_0)) = \mathcal{R}(\hat{E}_0),
\]
where \( \hat{E}_0 \subseteq \hat{L} \) is as in Section 27.1, and \( \mathcal{R}(\hat{E}_0) \) is the ideal in \( \hat{L} \) generated by \( \hat{E}_0 \).
This implies that there is a Lie algebra automorphism \( \hat{\sigma} \) of \( \hat{L} \), whose composition with the quotient mapping from \( L \) onto \( \hat{L} \) is the same as the composition of the quotient mapping with \( \hat{\sigma} \).

Of course,
\[
(27.6.7) \quad \hat{\sigma}(\tilde{E}_\alpha) = -\tilde{y}_\alpha, \quad \hat{\sigma}(\tilde{y}_\alpha) = -\tilde{x}_\alpha, \quad \text{and} \quad \hat{\sigma}(\tilde{h}_\alpha) = -\tilde{h}_\alpha
\]
for every \( \alpha \in \Delta \). This implies that
\[
(27.6.8) \quad \hat{\sigma}(\tilde{X}) = \tilde{X},
\]
which is the Lie subalgebra of \( \hat{L} \) generated by the \( \tilde{x}_\alpha \)'s, \( \alpha \in \Delta \), as before. In particular, \( \tilde{X} \) and \( \tilde{Y} \) are isomorphic as Lie algebras over \( k \), as in Exercise 1 on p101 of [13].

27.7 Some more relations in \( \hat{L} \)

Let us continue with the same notation and hypotheses as in the previous sections. Remember that \( n(\alpha, \beta) \in \mathbb{Z} \) for every \( \alpha, \beta \in \Delta \), with \( -3 \leq n(\alpha, \beta) \leq 0 \) when \( \alpha \neq \beta \). Put
\[
(27.7.1) \quad \tilde{x}_{\alpha, \beta} = (\text{ad}_{\tilde{x}_\alpha})^{-n(\beta, \alpha)+1}(\tilde{x}_\beta), \quad \tilde{y}_{\alpha, \beta} = (\text{ad}_{\tilde{y}_\alpha})^{-n(\beta, \alpha)+1}(\tilde{y}_\beta)
\]
for every \( \alpha, \beta \in \Delta \) with \( \alpha \neq \beta \), which are elements of \( \hat{L} \). We would like to show that
\[
(27.7.2) \quad \text{ad}_{\tilde{x}_\gamma}(\tilde{y}_{\alpha, \beta}) = 0
\]
for every \( \alpha, \beta, \gamma \in \Delta \) with \( \alpha \neq \beta \), as in the lemma on p99 of [13]. We also have that
\[
(27.7.3) \quad \text{ad}_{\tilde{y}_\gamma}(\tilde{x}_{\alpha, \beta}) = 0
\]
when \( \alpha \neq \beta \), using a similar argument, or the automorphism \( \hat{\sigma} \) on \( \hat{L} \) discussed in the previous section.

Suppose first that \( \alpha \neq \gamma \), so that \( [\tilde{x}_\gamma, \tilde{y}_\alpha] = 0 \), as in (27.1.10). This implies that \( \text{ad}_{\tilde{x}_\gamma} \) and \( \text{ad}_{\tilde{y}_\alpha} \) commute as linear mappings on \( \hat{L} \), as in Section 2.4. It follows that
\[
(27.7.4) \quad \text{ad}_{\tilde{x}_\gamma}(\tilde{y}_{\alpha, \beta}) = (\text{ad}_{\tilde{x}_\gamma})(\text{ad}_{\tilde{y}_\alpha})^{-n(\beta, \alpha)+1}(\tilde{x}_\beta)
\]
in this case. If \( \beta \neq \gamma \), then \( \text{ad}_{\tilde{x}_\gamma}(\tilde{y}_\beta) = 0 \), by (27.1.10) again, so that (27.7.2) holds. If \( \beta = \gamma \), then (27.7.4) reduces to
\[
(27.7.5) \quad \text{ad}_{\tilde{x}_\gamma}(\tilde{y}_{\alpha, \beta}) = (\text{ad}_{\tilde{y}_\alpha})^{-n(\beta, \alpha)+1}(\tilde{h}_\beta),
\]
27.8. SERRE’S THEOREM

by (27.1.10).

Note that
\[ \text{ad}_{\tilde{g}_\alpha} (\tilde{h}_\beta) = n(\alpha, \beta) \tilde{g}_\alpha, \]
by (27.1.12). If \( n(\alpha, \beta) = 0 \), then it follows that (27.7.2) holds, by (27.7.5). Otherwise, if \( n(\alpha, \beta) \neq 0 \), then \( n(\beta, \alpha) \neq 0 \), by (27.1.1). More precisely, \( n(\beta, \alpha) < 0 \), because \( \alpha \neq \beta \), so that
\[ n(\beta, \alpha) + 1 \geq 2. \]
This means that (27.7.2) follows from (27.7.5), because \( \text{ad}_{\tilde{g}_\alpha}(\tilde{g}_\alpha) = 0 \).

Suppose now that \( \alpha = \gamma \), so that we would like to show that
\[ \text{ad}_{\tilde{g}_\alpha}(\tilde{g}_\alpha) = 0 \]
when \( \alpha \neq \gamma \). Remember that \( \tilde{x}_\alpha, \tilde{y}_\alpha, \tilde{h}_\alpha \neq 0 \), as in (27.3.7) and (27.3.8). The linear span of \( \tilde{x}_\alpha, \tilde{y}_\alpha, \) and \( \tilde{h}_\alpha \) in \( \tilde{L} \) is a Lie subalgebra of \( \tilde{L} \), because of (27.1.10), (27.1.11), and (27.1.12). This subalgebra is isomorphic to \( sl(2, k) \) as a Lie algebra over \( k \), because \( n(\alpha, \alpha) = 2 \). Thus we may consider \( \tilde{L} \) as a module over \( sl(2, k) \), as a Lie algebra over \( k \), using \( \text{ad}_w \) when \( w \) is in the linear span of \( \tilde{x}_\alpha, \tilde{y}_\alpha, \) and \( \tilde{h}_\alpha \).

Note that \( \text{ad}_{\tilde{x}_\alpha}(\tilde{g}_\beta) = 0 \), as in (27.1.10), because \( \alpha \neq \beta \). We also have that
\[ \text{ad}_{\tilde{h}_\alpha}(\tilde{g}_\beta) = -n(\beta, \alpha) \tilde{g}_\beta, \]
as in (27.1.12). This means that \( \tilde{g}_\beta \) is a maximal or primitive vector of weight \( \lambda = -n(\beta, \alpha) \) in \( \tilde{L} \) as a module over the linear span of \( \tilde{x}_\alpha, \tilde{y}_\alpha, \) and \( \tilde{h}_\alpha \), as in Section 15.2. If \( j \) is a positive integer, then
\[ \text{ad}_{\tilde{x}_\alpha}((\text{ad}_{\tilde{g}_\alpha})^j(\tilde{g}_\beta)) = j(\lambda - j + 1)(\text{ad}_{\tilde{g}_\alpha})^j-1(\tilde{g}_\beta), \]
as in (15.3.4). The right side is equal to 0 when \( j = \lambda + 1 = -n(\beta, \alpha) + 1 \), which implies (27.7.8), as desired.

27.8 Serre’s theorem

Let us continue with the same notation and hypotheses as in the previous sections, Section 27.1 in particular. If \( \alpha, \beta \in \Delta \) and \( \alpha \neq \beta \), then put
\[ \tilde{x}_{\alpha, \beta} = (\text{ad}_{\tilde{x}_\alpha})^{-n(\beta, \alpha) + 1}(\tilde{x}_\beta), \quad \tilde{y}_{\alpha, \beta} = (\text{ad}_{\tilde{y}_\alpha})^{-n(\beta, \alpha) + 1}(\tilde{y}_\beta), \]
which are elements of \( \tilde{L} \). Of course, \( \text{ad}_{\tilde{x}_\alpha} \) and \( \text{ad}_{\tilde{y}_\alpha} \) refer to the adjoint representation on \( \tilde{L} \) here. Remember that \( n(\beta, \alpha) \leq 0 \) when \( \alpha \neq \beta \), so that \( -n(\beta, \alpha) + 1 \in \mathbb{Z}_+ \). Clearly the images of \( \tilde{x}_{\alpha, \beta} \) and \( \tilde{y}_{\alpha, \beta} \) in \( \tilde{L} \) under the natural quotient mapping are the same as \( \tilde{x}_{\alpha, \beta} \) and \( \tilde{y}_{\alpha, \beta} \) in (27.7.1), respectively.

Let \( \tilde{E}_1 \) be the subset of \( \tilde{L} \) consisting of \( \tilde{x}_{\alpha, \beta} \) and \( \tilde{y}_{\alpha, \beta} \) for \( \alpha, \beta \in \Delta \) with \( \alpha \neq \beta \). Remember that \( \tilde{E}_0 \subseteq \tilde{L} \) was defined in Section 27.1, and let \( \mathcal{R}(\tilde{E}_0 \cup \tilde{E}_1) \) be the ideal in \( \tilde{L} \) generated by \( \tilde{E}_0 \cup \tilde{E}_1 \). Thus the quotient
\[ \tilde{L}/\mathcal{R}(\tilde{E}_0 \cup \tilde{E}_1) \]
is a Lie algebra over $k$. A famous theorem of Serre states that this is a finite-dimensional semisimple Lie algebra, as on p52 of [23], and p99 of [13]. More precisely, the linear span of the images of the $\tilde{h}_\alpha$’s, $\alpha \in \Delta$, in (27.8.2) is a Cartan subalgebra of this Lie algebra, which corresponds to the given root system $\Phi$.

Let $\tilde{E}_1$ be the subset of $\tilde{L}$ consisting of $\tilde{x}_{\alpha,\beta}$ and $\tilde{y}_{\alpha,\beta}$ for $\alpha, \beta \in \Delta$ with $\alpha \neq \beta$, which is the same as the image of $\tilde{E}_1$ under the natural quotient mapping from $\tilde{L}$ onto $\tilde{L}$. Also let $\mathcal{R}(\tilde{E}_1) = \mathcal{R}_{\tilde{E}}(\tilde{E}_1)$ be the ideal in $\tilde{L}$ generated by $\tilde{E}_1$. Consider the quotient
\[(27.8.3) \quad L = L(k) = \tilde{L}/\mathcal{R}(\tilde{E}_1),\]
which is a Lie algebra over $k$. There is a natural Lie algebra homomorphism from $\tilde{L}$ onto $L$, which is the composition of the quotient mappings from $\tilde{L}$ onto $L$ and from $\tilde{L}$ onto $\tilde{L}$. The kernel of this homomorphism is $\mathcal{R}(\tilde{E}_0 \cup \tilde{E}_1)$, which leads to a natural Lie algebra isomorphism from (27.8.2) onto $L$.

Let $x_\alpha$, $y_\alpha$, and $h_\alpha$ be the images of $\tilde{x}_\alpha$, $\tilde{y}_\alpha$, and $\tilde{h}_\alpha$, respectively, under the natural quotient mapping from $L$ onto $L$ for each $\alpha \in \Delta$. Of course, these correspond to the images of $\tilde{x}_\alpha$, $\tilde{y}_\alpha$, and $\tilde{h}_\alpha$, respectively, in (27.8.2) for every $\alpha \in \Delta$. If $\alpha, \beta \in \Delta$, then
\[
[27.8.4] \quad [h_\alpha, h_\beta] = 0,
\]
\[
[27.8.5] \quad [x_\alpha, y_\beta] = \delta_{\alpha,\beta} h_\alpha,
\]
\[
[27.8.6] \quad [h_\alpha, x_\beta] = n(\beta, \alpha) x_\beta,
\]
\[
[27.8.7] \quad [h_\alpha, y_\beta] = -n(\beta, \alpha) y_\beta
\]
in $L$, because of the analogous properties of $\tilde{x}_\alpha$, $\tilde{y}_\alpha$, and $\tilde{h}_\alpha$ in $\tilde{L}$, as in Section 27.1. If $\alpha \neq \beta$, then we also have that
\[
[27.8.8] \quad (\text{ad}_{x_\alpha})^{-n(\beta, \alpha)+1}(x_\beta) = (\text{ad}_{y_\alpha})^{-n(\beta, \alpha)+1}(y_\beta) = 0
\]
in $L$, by definition of $\tilde{E}_1$.

The proof of Serre’s theorem will be discussed in the next sections, and continuing in the next chapter.

### 27.9 Ideals $\mathcal{I}$, $\mathcal{J}$

We continue with the discussion of Serre’s theorem, from the previous section.

Remember that $\tilde{X}$, $\tilde{Y}$ are the Lie subalgebras of $\tilde{L}$ generated by the $\tilde{x}_\alpha$’s, $\tilde{y}_\alpha$’s, respectively, with $\alpha \in \Delta$, as in Section 27.3. Let $\tilde{E}_{1,1}$ be the subset of $\tilde{L}$ consisting of $\tilde{x}_{\alpha,\beta}$ for $\alpha, \beta \in \Delta$ with $\alpha \neq \beta$, and let $\tilde{E}_{1,2}$ be the set of $\tilde{y}_{\alpha,\beta}$ for $\alpha, \beta \in \Delta$ with $\alpha \neq \beta$. Thus
\[
[27.9.1] \quad \tilde{E}_1 = \tilde{E}_{1,1} \cup \tilde{E}_{1,2}, \: \tilde{E}_{1,1} \subseteq \tilde{X}, \: \text{and} \: \tilde{E}_{1,2} \subseteq \tilde{Y}.
\]

Let $\mathcal{I}$ be the ideal in $\tilde{X}$ generated by $\tilde{E}_{1,1}$, and let $\mathcal{J}$ be the ideal in $\tilde{Y}$ generated by $\tilde{E}_{1,2}$. Note that
\[
[27.9.2] \quad \mathcal{I}, \mathcal{J} \subseteq \mathcal{R}(\tilde{E}_1).
\]
27.10. **Subalgebras $\mathcal{X}$, $\mathcal{H}$, and $\mathcal{Y}$ of $L$**

We would like to show that

$\mathcal{I}, \mathcal{J}$ are ideals in $\tilde{L}$, as in Step (1) on p99 of [13], and (a) on p53 of [23]. We shall do this for $\mathcal{J}$, since the argument for $\mathcal{I}$ is analogous, as in [13]. Observe that

\begin{equation}
[\tilde{h}_\gamma, \tilde{y}_{\alpha, \beta}] = (-n(\beta, \gamma) + n(\beta, \alpha) - 1) n(\alpha, \gamma) \tilde{y}_{\alpha, \beta}
\end{equation}

for every $\alpha, \beta, \gamma \in \Delta$ with $\alpha \neq \beta$, by (27.4.3) and the definition (27.7.1) of $\tilde{y}_{\alpha, \beta}$. It is easy to see that

\begin{equation}
ad_{\tilde{\gamma}}(\tilde{\mathcal{Y}}) \subseteq \tilde{\mathcal{Y}}
\end{equation}

for every $\gamma \in \Delta$, using (27.1.12) and the Jacobi identity; a more precise version of this is given by (27.4.3). It follows from (27.9.4) and (27.9.5) that

\begin{equation}
ad_{\tilde{\gamma}}(\tilde{\mathcal{J}}) \subseteq \tilde{\mathcal{J}}
\end{equation}

for every $\gamma \in \Delta$, using the Jacobi identity again.

Remember that $\mathcal{H}$ is the linear span of the $\tilde{h}_\alpha$’s, $\alpha \in \Delta$, in $\tilde{L}$, as in Section 27.3. Observe that

\begin{equation}
ad_{\tilde{\gamma}}(\tilde{\mathcal{Y}}) \subseteq \tilde{\mathcal{Y}} + \tilde{\mathcal{H}}
\end{equation}

for every $\gamma \in \Delta$, by (27.1.10) and the earlier remarks about (27.4.6). One can check that

\begin{equation}
ad_{\tilde{\gamma}}(\tilde{\mathcal{J}}) \subseteq \tilde{\mathcal{J}}
\end{equation}

for every $\gamma \in \Delta$, using (27.7.2), (27.9.6), (27.9.7), and the Jacobi identity. This implies that $\mathcal{J}$ is an ideal in $\tilde{L}$, because $\tilde{L}$ is generated as a Lie algebra over $k$ by the $\tilde{x}_\gamma$’s and $\tilde{y}_\gamma$’s, $\gamma \in \Delta$. More precisely, this is the same as saying that the normalizer of $\mathcal{J}$ in $\tilde{L}$ is $\tilde{L}$, and the normalizer of $\mathcal{J}$ in $\tilde{L}$ is automatically a Lie subalgebra of $\tilde{L}$, by the Jacobi identity.

Step (2) on p99 of [13] states that

\begin{equation}
\mathcal{I} + \mathcal{J} = \mathcal{R}(\tilde{E}_1),
\end{equation}

which corresponds to a remark just after (a) on p53 of [23]. Of course,

\begin{equation}
\mathcal{I} + \mathcal{J} \subseteq \mathcal{R}(\tilde{E}_1),
\end{equation}

by (27.9.2). We also have that $\tilde{E}_1 \subseteq \mathcal{I} + \mathcal{J}$, by construction. Note that $\mathcal{I} + \mathcal{J}$ is an ideal in $\tilde{L}$, by (27.9.3). This implies that $\mathcal{R}(\tilde{E}_1) \subseteq \mathcal{I} + \mathcal{J}$, as desired.

27.10 **Subalgebras $\mathcal{X}$, $\mathcal{H}$, and $\mathcal{Y}$ of $L$**

Let us continue with the same notation and hypotheses as in the previous sections. Let $\mathcal{X}$, $\mathcal{H}$, and $\mathcal{Y}$ be the images of $\tilde{\mathcal{X}}$, $\tilde{\mathcal{H}}$, and $\tilde{\mathcal{Y}}$, respectively, under the
natural quotient mapping from $\tilde{L}$ onto $L$. Thus $\mathcal{X}$, $\mathcal{H}$, and $\mathcal{Y}$ are Lie subalgebras of $L$, and $\mathcal{H}$ is commutative as a Lie algebra over $k$, because of the analogous properties of $\tilde{\mathcal{X}}$, $\tilde{\mathcal{H}}$, and $\tilde{\mathcal{Y}}$ in $\tilde{L}$. More precisely, $\mathcal{X}$ is the Lie subalgebra of $L$ generated by the $x_\alpha$'s, $\alpha \in \Delta$, and $\mathcal{Y}$ is the Lie subalgebra of $L$ generated by the $y_\alpha$'s, $\alpha \in \Delta$. Similarly, $\mathcal{H}$ is the same as the linear span of the $h_\alpha$'s, $\alpha \in \Delta$, in $L$.

Remember that $\tilde{L}$ corresponds to the direct sum of $\tilde{\mathcal{X}}$, $\tilde{\mathcal{H}}$, and $\tilde{\mathcal{Y}}$ as a vector space over $k$, as in (27.3.2). Similarly, $L$ corresponds to the direct sum of $\mathcal{X}$, $\mathcal{H}$, and $\mathcal{Y}$, as a vector space over $k$.

This follows from (27.9.9), and the analogous statement for $\tilde{L}$. This is part of Step (3) on p100 of [13], and of (b) on p53 of [23].

The kernels of the restrictions of the natural quotient mapping from $\tilde{L}$ onto $L$ to $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{Y}}$ are $I$ and $J$, respectively, because of (27.9.9). Thus we can identify $\tilde{\mathcal{X}}$, $\tilde{\mathcal{Y}}$ with the quotients

$$\tilde{\mathcal{X}}/I, \quad \tilde{\mathcal{Y}}/J,$$

respectively. Similarly, the restriction of the natural quotient mapping from $\tilde{L}$ onto $L$ to $\tilde{\mathcal{H}}$ is injective, by (27.9.9). This permits us to identify $\mathcal{H}$ with $\mathcal{H}$. This is also part of Step (3) on p100 of [13], and (b) on p53 of [23].

Remember that the $\tilde{h}_\alpha$'s, $\alpha \in \Delta$, form a basis for $\tilde{\mathcal{H}}$ as a vector space over $k$, as in (27.3.1). This means that the $h_\alpha$'s, $\alpha \in \Delta$, form a basis for $\mathcal{H}$ as a vector space over $k$. In particular, $\tilde{h}_\alpha \neq 0$ for each $\alpha \in \Delta$, as in (27.3.7), so that $h_\alpha \neq 0$ for every $\alpha \in \Delta$ too. This implies that $x_\alpha, y_\alpha \neq 0$ for every $\alpha \in \Delta$, because $[x_\alpha, y_\alpha] = h_\alpha$, as in (27.8.5).

Using this, one can verify that

$$\text{the set of } x_\beta \text{'s, } y_\beta \text{'s, and } h_\beta \text{'s, } \beta \in \Delta,$$

is linearly independent in $L$, as in (27.3.10). More precisely, the $x_\beta$'s, $y_\beta$'s, and $h_\beta$'s are simultaneous eigenvectors for $\text{ad}_{h_\alpha}$ on $L$, $\alpha \in \Delta$, by (27.8.4), (27.8.6), and (27.8.7). To get linear independence, one also uses the fact that the eigenvalues for the $x_\beta$'s and $y_\beta$'s correspond to distinct functions of $\alpha$ on $\Delta$ that are not identically 0, as before. This permits one to reduce to the linear independence of the $h_\beta$'s, $\beta \in \Delta$. Equivalently, the restriction of the natural quotient mapping from $\tilde{L}$ onto $L$ to the linear subspace of $\tilde{L}$ spanned by the $\tilde{x}_\beta$'s, $\tilde{y}_\beta$'s, and $\tilde{h}_\beta$'s, $\beta \in \Delta$, is injective, as in Step (4) on p100 of [13].

### 27.11 Eigenvectors in $L$

We continue with the same notation and hypotheses as in the previous sections.
Let $H'$ be the dual of $H$, as a vector space over $k$, as usual, and put
\[(27.11.1) \quad L_\mu = \{ u \in L : [h, u] = \mu(h) u \text{ for every } h \in H \}\]
for each $\mu \in H'$. The elements of $L_\mu$ are said to have weight $\mu$. If $\mu \in H'$, then let $\tilde{\mu}$ be the corresponding linear functional on $\bar{H}$, which is the composition of $\mu$ with the restriction to $\bar{H}$ of the natural quotient mapping from $\bar{L}$ onto $L$. Observe that
\[(27.11.2) \quad \text{the natural quotient mapping from } \bar{L} \text{ onto } L \text{ maps } \bar{L}_\mu \text{ into } L_\mu, \]
where $\bar{L}_\mu$ is as in (27.5.1). Of course, $H \subseteq L_0$, by (27.8.4).

If $\beta \in \Delta$, then there is a unique $\nu_\beta \in H'$ such that
\[(27.11.3) \quad \nu_\beta(h_\alpha) = n(\beta, \alpha) \]
for every $\alpha \in \Delta$, because the $h_\alpha$’s form a basis for $H$. The composition of $\nu_\beta$ with the restriction of the natural quotient mapping from $\bar{L}$ onto $L$ to $H$ is the same as $\tilde{\nu}_\beta \in \bar{H}'$, defined in (27.5.2). It is easy to see that the $\nu_\beta$’s, $\beta \in \Delta$, form a basis for $H'$, as a vector space over $k$, because of the analogous statement for the $\tilde{\nu}_\beta$’s in Section 27.5.

Every element of $\mathcal{X}$ can be expressed as the sum of finitely many elements of the subspaces $L_\mu$, where $\mu \in H'$ can be expressed as a nontrivial sum of $\nu_\beta$’s. This follows from (27.11.2) and the analogous statement for $\bar{X}$ in Section 27.5. We also have that every element of $\mathcal{Y}$ can be expressed as the sum of finitely many elements of $L_\mu$’s, where $\mu \in H'$ can be expressed as a nontrivial sum of $-\nu_\beta$’s.

This implies that $\mathcal{X}$ and $\mathcal{Y}$ correspond to the direct sums of their intersections with the appropriate subspaces $L_\mu$, as vector spaces over $k$, as in Section 27.5.

As before, one can use this to check that
\[(27.11.4) \quad H = L_0. \]

Similarly, $\mathcal{X}$ corresponds to the direct sum of the subspaces $L_\mu$, where $\mu \in H'$ can be expressed as a nontrivial sum of $\nu_\beta$’s, and $\mathcal{Y}$ corresponds to the direct sum of the subspaces $L_\mu$, where $\mu \in H'$ can be expressed as a nontrivial sum of $-\nu_\beta$’s, as vector spaces over $k$. This uses the hypothesis that $k$ have characteristic 0, so that nontrivial sums of $\nu_\beta$’s or of $-\nu_\beta$’s cannot be equal to 0, or to each other. If $\mu \in H'$ and $L_\mu \neq \{0\}$, then one can verify that $\mu = 0$, or $\mu$ is a nontrivial sum of $\nu_\beta$’s, or $\mu$ is a nontrivial sum of $-\nu_\beta$’s.

More precisely, every element of $\mathcal{X}$ can be expressed as the sum of the images under the natural quotient mapping from $L$ onto $L$ of finitely many elements of $\bar{L}_\mu$, where $\tilde{\mu}$ is a nontrivial sum of $\tilde{\nu}_\beta$’s, and every element of $\mathcal{Y}$ can be expressed as the sum of the images of finitely many elements of $\bar{L}_\mu$, where $\tilde{\mu}$ is a nontrivial sum of $-\tilde{\nu}_\beta$’s. Let $\mu \in H'$ be given, and let $\tilde{\mu}$ be the corresponding linear functional on $\bar{H}'$ again, as in (27.11.2). One can check that
\[(27.11.5) \quad \text{the natural quotient mapping from } \bar{L} \text{ onto } L \text{ maps } \bar{L}_\mu \text{ onto } L_\mu. \]
This follows from (27.11.4) when $\mu = 0$. Otherwise, one can use the previous remarks about $X$, $Y$.

Using (27.11.5), we get that $L_\mu$ is finite-dimensional as a vector space over $k$ for every $\mu \in \mathcal{H}'$, because of the analogous statement for $L_{-\mu}$ in Section 27.5. This corresponds to Step (5) on p100 of [13], and some remarks on p53 of [23].

### 27.12 Locally nilpotent linear mappings

Let $k$ be a field, and let $V$ be a vector space over $k$. A linear mapping $T$ from $V$ into itself is said to be **locally nilpotent** if for each $v \in V$ there is a positive integer $l$ such that $T^l(v) = 0$. If $W$ is a finite-dimensional linear subspace of $V$, then it follows that there is a positive integer $r$ such that $T^r(W) = \{0\}$. In this case, the linear subspace $W_1$ of $V$ spanned by $W, T(W), \ldots, T^r(W)$ is finite-dimensional and satisfies

$$
(27.12.1) \quad T(W_1) \subseteq W_1.
$$

If $T_1, T_2$ are commuting locally nilpotent linear mappings from $V$ into itself, then it is easy to see that $T_1 + T_2$ is locally nilpotent on $V$ as well.

Suppose now that $k$ has characteristic 0, and let $T$ be a locally nilpotent linear mapping from $V$ into itself. If $v \in V$, then put

$$
(27.12.2) \quad (\exp T)(v) = \sum_{j=0}^{\infty} \frac{1}{j!} T^j(v),
$$

where the sum on the right reduces to a finite sum in $V$. It is easy to see that this defines a linear mapping from $V$ into itself.

Let $T_1, T_2$ be commuting locally nilpotent linear mappings from $V$ into itself again. If $v \in V$, then

$$
(27.12.3) \quad T_1^j(T_2^l(v)) = T_2^l(T_1^j(v)) = 0
$$

when $j$ or $l$ is sufficiently large. One can check that

$$
(27.12.4) \quad (\exp(T_1 + T_2))(v) = (\exp T_1)(\exp T_2)(v),
$$

using the binomial theorem in the usual way. If $T$ is a locally nilpotent linear mapping from $V$ into itself, then one can apply the previous statement to $T$, $-T$ to get that $\exp T$ is invertible on $V$, with inverse equal to $\exp(-T)$. This is related to some remarks just before the theorem on p99 of [13].

Let $\mathcal{A}$ be an algebra over $k$ in the strict sense, and let $\delta$ be a derivation on $\mathcal{A}$. If $\delta$ is locally nilpotent on $\mathcal{A}$, then $\exp \delta$ is an algebra automorphism of $\mathcal{A}$, as in Section 14.11.

Let us now return to the discussion of Serre’s theorem, as in the previous sections. We would like to show that for each $\alpha \in \Delta$, $\text{ad}_{x \alpha}$ and $\text{ad}_{y \alpha}$ are locally nilpotent as linear mappings from $L$ into itself, as in Step (6) on p100 of [13], and (c) on p53 of [23]. Put

$$
(27.12.5) \quad V_\alpha = \{ u \in L : \text{ad}^l_{x \alpha}(u) = 0 \text{ for some } l \in \mathbb{Z}_+ \},
$$
27.13. SOME PROPERTIES OF $\theta_\alpha$

which is a linear subspace of $L$. One can check that $V_\alpha$ is a Lie subalgebra of $L$, using the fact that $\text{ad}_{x_\alpha}$ is a derivation on $L$. Observe that $x_\beta \in V_\alpha$ for every $\beta \in \Delta$, by (27.8.8). One can verify that $y_\beta \in V_\alpha$ for every $\beta \in \Delta$, using (27.8.5) and (27.8.6). It follows that $L \subseteq V_\alpha$, because $L$ is generated as a Lie algebra over $k$ by the $x_\beta$'s and $y_\beta$'s, $\beta \in \Delta$. This means that $\text{ad}_{x_\alpha}$ is locally nilpotent on $L$, and the argument for $\text{ad}_{y_\alpha}$ is analogous.

Let $\alpha \in \Delta$ be given, and note that $-\text{ad}_{y_\alpha}$ is locally nilpotent on $L$. Thus $\exp \text{ad}_{x_\alpha}$ and $\exp(-\text{ad}_{y_\alpha})$ are Lie algebra automorphisms of $L$, because $\text{ad}_{x_\alpha}$, $-\text{ad}_{y_\alpha}$ are derivations on $L$, as before. It follows that

$$\theta_\alpha = \exp \text{ad}_{x_\alpha} \circ \exp(-\text{ad}_{y_\alpha}) \circ \exp \text{ad}_{x_\alpha}$$

is a Lie algebra automorphism of $L$ too. This corresponds to Step (7) on p100 of [13], and to part of (d) on p54 of [23].

If $u \in L$ satisfies $[x_\alpha, u] = [y_\alpha, u] = 0$, then

$$\exp \text{ad}_{x_\alpha}(u) = \exp(-\text{ad}_{y_\alpha})(u) = u,$$

and hence

$$\theta_\alpha(u) = u.$$

Remember that $n(\alpha, \alpha) = 2$, as in Section 27.1. Using the same type of argument as in Section 23.7, we get that

$$\theta_\alpha(x_\alpha) = -y_\alpha, \quad \theta_\alpha(y_\alpha) = -x_\alpha, \quad \theta_\alpha(h_\alpha) = -h_\alpha.$$

27.13 Some properties of $\theta_\alpha$

Let us continue with the same notation and hypotheses as in the previous sections. In particular, let $\alpha \in \Delta$ be given again, and let $\theta_\alpha$ be the Lie algebra automorphism of $L$ defined in (27.12.6). If $h \in H$, then $h$ can be expressed in a unique way as

$$h = \sum_{\beta \in \Delta} c_\beta h_\beta,$$

with $c_\beta \in k$ for each $\beta \in \Delta$. In this case,

$$[h, x_\alpha] = \sum_{\beta \in \Delta} c_\beta [h_\beta, x_\alpha] = \sum_{\beta \in \Delta} c_\beta n(\alpha, \beta) x_\alpha,$$

$$[h, y_\alpha] = \sum_{\beta \in \Delta} c_\beta [h_\beta, y_\alpha] = -\sum_{\beta \in \Delta} c_\beta n(\alpha, \beta) y_\alpha,$$

by (27.8.6) and (27.8.7). If

$$\sum_{\beta \in \Delta} c_\beta n(\alpha, \beta) = 0,$$
then it follows that \([h, x_\alpha] = [h, y_\alpha] = 0\), so that

\[(27.13.5) \quad \theta_\alpha(h) = h,\]

as in (27.12.8).

If \(h = h_\alpha\), then (27.13.4) does not hold, because \(k\) has characteristic 0, so that \(n(\alpha, \alpha) = 2\) is not zero in \(k\). It follows that

\[(27.13.6) \quad \mathcal{H} \text{ is spanned by } h_\alpha \text{ and the } h \in \mathcal{H} \text{ that satisfy } (27.13.4),\]

as a vector space over \(k\). This implies that

\[(27.13.7) \quad \theta_\alpha(\mathcal{H}) = \mathcal{H},\]

by (27.12.9) and (27.13.5). Note that

\[(27.13.8) \quad \theta_\alpha(\theta_\alpha(h)) = h\]

for every \(h \in \mathcal{H}\).

If \(h \in \mathcal{H}\) is as in (27.13.1), then put

\[(27.13.9) \quad \nu_\alpha(h) = \sum_{\beta \in \Delta} c_\beta n(\alpha, \beta).\]

This defines a linear functional on \(\mathcal{H}\), which is the same as the one in (27.11.3), with \(\alpha\) and \(\beta\) interchanged. Observe that \(\nu_\alpha(h_\alpha) = n(\alpha, \alpha) = 2\), and that

\[(27.13.10) \quad \theta_\alpha(h) = h - \nu_\alpha(h) h_\alpha\]

for every \(h \in \mathcal{H}\), because of (27.12.9) and (27.13.5).

Let \(\mu \in \mathcal{H}'\) be given, and put

\[(27.13.11) \quad \zeta_{\mu, \alpha}(h) = \mu(\theta_\alpha(h))\]

for every \(h \in \mathcal{H}\), which defines a linear functional on \(\mathcal{H}\) too. Clearly \(h \in \mathcal{H}\) and \(u \in L\) satisfy

\[(27.13.12) \quad [h, u] = \zeta_{\mu, \alpha}(h) u\]

if and only if

\[(27.13.13) \quad [\theta_\alpha(h), \theta_\alpha(u)] = \mu(\theta_\alpha(h)) \theta_\alpha(u),\]

because \(\theta_\alpha\) is a Lie algebra automorphism of \(L\). This implies that

\[(27.13.14) \quad \theta_\alpha(L_{\zeta_{\mu, \alpha}}) = L_{\mu},\]

where \(L_{\mu}, L_{\zeta_{\mu, \alpha}}\) are as in (27.11.1). This corresponds to part of Step (8) on p100 of [13], and of (d) on p54 of [23].

Equivalently,

\[(27.13.15) \quad \zeta_{\mu, \alpha}(h) = \mu(h) - \nu_\alpha(h) \mu(h_\alpha)\]
27.14 Roots and $\mathcal{H}'$

for every $h \in \mathcal{H}$, by (27.13.10). This means that

\begin{equation}
(27.13.16) \quad \zeta_{\mu, \alpha} = \mu - \mu(h_\alpha) \nu_\alpha,
\end{equation}

as elements of $\mathcal{H}'$. Let $\gamma \in \Delta$ be given, and let $\nu_\gamma \in \mathcal{H}'$ be as in (27.11.3) and (27.13.9), so that

\begin{equation}
(27.13.17) \quad \nu_\gamma(h_\beta) = n(\gamma, \beta)
\end{equation}

for every $\beta \in \Delta$. If we take $\mu = \nu_\gamma$ in (27.13.16), then we get that

\begin{equation}
(27.13.18) \quad \zeta_{\nu_\gamma, \alpha} = \nu_\gamma - \nu_\gamma(h_\alpha) \nu_\alpha = \nu_\gamma - n(\gamma, \alpha) \nu_\alpha.
\end{equation}

This means that

\begin{equation}
(27.13.19) \quad \zeta_{\nu_\gamma, \alpha} = \nu_\gamma - \lambda_\alpha(\gamma) \nu_\alpha,
\end{equation}

where $\lambda_\alpha(\gamma) \in \mathbb{Z}$ is as in Section 27.1.

27.14 Roots and $\mathcal{H}'$

We continue with the discussion of Serre’s theorem, as in the previous sections. Remember from Section 27.1 that $V$ is a vector space over the real numbers of positive finite dimension, $\Phi$ is a reduced root system in $V$, and that $\Delta$ is a base for $\Phi$. Let $\Theta_\Phi$ be the subgroup of $V$, as a commutative group with respect to addition, generated by $\Phi$. This is the same as the subgroup of $V$ generated by $\Delta$, which consists of linear combinations of elements of $\Delta$ with coefficients in $\mathbb{Z}$.

If $\alpha \in \Phi$, then $\sigma_\alpha$ is the symmetry on $V$ with vector $\alpha$ that maps $\Phi$ onto itself, as before. This can be expressed as $\sigma_\alpha(v) = v - \lambda_\alpha(v) \alpha$, where $\lambda_\alpha$ is a linear functional on $V$ such that $\lambda_\alpha(\alpha) = 2$ and $\lambda_\alpha(\beta) \in \mathbb{Z}$ for every $\beta \in \Phi$. Observe that $\sigma_\alpha$ maps $\Theta_\Phi$ onto itself, and that $\lambda_\alpha$ maps $\Theta_\Phi$ into $\mathbb{Z}$. If $\sigma$ is in the Weyl group $W$ of $\Phi$, then it follows that

\begin{equation}
(27.14.1) \quad \sigma(\Theta_\Phi) = \Theta_\Phi.
\end{equation}

If $\gamma \in \Theta_\Phi$, then let $\nu_\gamma$ be the linear functional on $\mathcal{H}$ such that

\begin{equation}
(27.14.2) \quad \nu_\gamma(h_\beta) = \lambda_\beta(\gamma)
\end{equation}

for every $\beta \in \Delta$. This is equivalent to (27.13.17) when $\gamma \in \Delta$, by (27.1.1). Note that

\begin{equation}
(27.14.3) \quad \gamma \mapsto \nu_\gamma
\end{equation}

is a homomorphism from $\Theta_\Phi$ into $\mathcal{H}'$, as commutative groups with respect to addition. Remember that the $\nu_\gamma$’s, $\gamma \in \Delta$, form a basis for $\mathcal{H}'$ as a vector space over $k$, as in Section 27.11. This implies that (27.14.3) is injective on $\Theta_\Phi$, because $k$ has characteristic 0.

Let $\alpha \in \Delta$ be given, and let $\theta_\alpha$ be the Lie algebra automorphism of $L$ defined in Section 27.12. If $\gamma \in \Theta_\Phi$, then $\zeta_{\nu_\gamma, \alpha} \in \mathcal{H}'$ can be defined as in (27.13.11). Using (27.13.16), we get that

\begin{equation}
(27.14.4) \quad \zeta_{\nu_\gamma, \alpha} = \nu_\gamma - \nu_\gamma(h_\alpha) \nu_\alpha = \nu_\gamma - \lambda_\alpha(\gamma) \nu_\alpha.
\end{equation}
Equivalently, this means that
\[(27.14.5) \quad \zeta_{\nu, \alpha} = \nu_{\sigma, \gamma}.
\]
It follows that
\[(27.14.6) \quad \theta_{\alpha}(L_{\nu_{\sigma, \gamma}}) = L_{\nu_{\gamma}},
\]
by taking \(\mu = \nu_{\gamma}\) in (27.13.14).

Remember that \(L_{\mu}\) is finite-dimensional as a vector space over \(k\) for every \(\mu \in \mathcal{H}'\), as in Section 27.11. Using (27.14.6), we get that
\[(27.14.7) \quad \dim L_{\nu_{\alpha}} = \dim L_{\nu_{\gamma}}
\]
for every \(\alpha \in \Delta\) and \(\gamma \in \Theta_{\Phi}\). This implies that
\[(27.14.8) \quad \dim L_{\nu_{\sigma}} = \dim L_{\nu_{\gamma}}
\]
for every \(\sigma \in W\) and \(\gamma \in \Theta_{\Phi}\). More precisely, this uses the fact that \(W\) is generated by \(\sigma_{\alpha}\) with \(\alpha \in \Delta\), as in Section 19.14. This corresponds to Step (8) on p100 of [13], and (d) on p54 of [23].

### 27.15 Some dimensions of eigenspaces

We continue with the same notation and hypotheses as in the previous sections. Let \(\beta \in \Delta\) be given, and let \(\tilde{\nu}_{\beta}\) be the linear functional on \(\mathcal{H}\) such that \(\tilde{\nu}_{\beta}(h_{\alpha}) = n(\beta, \alpha)\) for every \(\alpha \in \Delta\), as in (27.5.2). Remember that \(\tilde{L}_{\nu_{\beta}}\) consists of the \(\tilde{u} \in \tilde{L}\) such that \([\tilde{h}, \tilde{u}] = \tilde{\nu}_{\beta}(\tilde{h}) \tilde{u}\) for every \(\tilde{h} \in \tilde{H}\), as in (27.5.1). Equivalently, this means that \([\tilde{h}_{\alpha}, \tilde{u}] = n(\beta, \alpha) \tilde{u}\) for every \(\alpha \in \Delta\). In particular, \(\tilde{x}_{\beta} \in \tilde{L}_{\nu_{\beta}}\), by (27.11.1).

In fact, \(\tilde{L}_{\nu_{\beta}}\) is spanned by \(\tilde{x}_{\beta}\). Remember that \(\tilde{L}\) is spanned by \(\tilde{X}, \tilde{H},\) and \(\tilde{Y}\), and more precisely by elements of \(\tilde{L}_{\nu_{\beta}}\) for suitable \(\tilde{\mu} \in \tilde{H}'\). Only multiples of \(\tilde{x}_{\beta}\) by elements of \(\mathcal{K}\) were needed when \(\tilde{\mu} = \tilde{\nu}_{\beta}\). This implies that \(\tilde{L}_{\nu_{\beta}}\) consists of only multiples of \(\tilde{x}_{\beta}\).

Similarly, if \(r \in \mathbb{Z}\) and \(r \neq 0, \pm 1\), then \(\tilde{L}_{\nu_{\beta}}\) is \(\{0\}\). Indeed, If \(|r| \geq 2\), then the only elements of \(\tilde{L}_{\nu_{\beta}}\) that would have been needed before were multiples of \([\tilde{y}_{\beta_1}, \ldots, \tilde{y}_{\beta_{|r|}}]\) with \(\beta_l = \beta\) for \(l = 1, \ldots, |r|\), as in Section 27.4. Of course, \([\tilde{x}_{\beta_1}, \ldots, \tilde{x}_{\beta_{|r|}}]\) and \([\tilde{y}_{\beta_1}, \ldots, \tilde{y}_{\beta_{|r|}}]\) are equal to \(0\) in this case. It follows that \(0\) is the only element of \(\tilde{L}_{\nu_{\beta}}\) when \(|r| \geq 2\).

Let \(\nu_{\beta}\) be the linear functional on \(\mathcal{H}\) such that \(\nu_{\beta}(h_{\alpha}) = n(\beta, \alpha)\) for every \(\alpha \in \Delta\), as in (27.11.3). Note that \(\tilde{\nu}_{\beta}\) is the same as the composition of \(\nu_{\beta}\) with the restriction to \(\tilde{H}\) of the natural quotient mapping from \(\tilde{L}\) onto \(L\). If \(r \in \mathbb{Z}\), then the natural quotient mapping from \(\tilde{L}\) onto \(L\) maps \(\tilde{L}_{\nu_{\beta}}\) onto \(L_{\nu_{\beta}}\), as in (27.11.5). This implies that
\[(27.15.1) \quad L_{\nu_{\beta}} = \{0\} \text{ when } |r| \geq 2.
\]
27.15. SOME DIMENSIONS OF EIGENSPACES

by the remarks in the preceding paragraph.

If we take \( r = 1 \), then we get that \( L_{\nu_\beta} \) is spanned by \( x_\beta \), because of the analogous statement for \( \tilde{L}_{\nu_\beta} \). Remember that \( x_\beta \neq 0 \), as in Section 27.10. Thus

\[(27.15.2)\quad \dim L_{\nu_\beta} = 1.\]

This corresponds to Step (9) on p100 of [13], and (e) on p54 of [23].

Let \( \Theta_\Phi \) be as in the previous section, and if \( \gamma \in \Theta_\Phi \), then let \( \nu_\gamma \in \mathcal{H}' \) be as in (27.14.2). Remember that this is equivalent to the previous definition when \( \gamma \in \Delta \). If \( \gamma \in \Phi \), then there is an element \( \sigma \) of the Weyl group of \( \Phi \) such that \( \sigma(\gamma) \in \Delta \), as in Section 19.14. It follows that

\[(27.15.3)\quad \dim L_{\nu_\gamma} = 1,\]

by (27.14.8) and (27.15.2).

If \( r \in \mathbb{Z} \), then \( r\gamma \in \Theta_\Phi \), and \( r\nu_\gamma = \nu_{r\gamma} \). Similarly, \( \nu_{\sigma(\gamma)} = \nu_{\sigma(\gamma)} = r\nu_{\sigma(\gamma)} \).

If \( |r| \geq 2 \), then we get that

\[(27.15.4)\quad \dim L_{r\nu_\gamma} = 0,\]

using (27.14.8) and (27.15.1). This corresponds to Step (10) on p100 of [13], and (f) on p54 of [23].
Chapter 28

Some more Lie algebras, 2

28.1 Linear combinations of roots

Let us continue with the discussion of Serre’s theorem, as in the previous chapter. We first need an auxiliary result about root systems. Remember that $(\cdot, \cdot)$ is an inner product on $V$ that is invariant under the Weyl group of $\Phi$, as in Section 27.1. If $w \in V$, then let $V'_w$ be the linear subspace of the dual $V'$ of $V$ consisting of linear functionals on $V$ that are equal to 0 at $w$, which is a hyperplane in $V'$ when $w \neq 0$. Of course, one can use $(\cdot, \cdot)$ to identify $V'$ with $V$, and $V'_w$ with the set of vectors in $V$ that are orthogonal to $w$.

Suppose that $\gamma \in V$ is not a multiple of any element of $\Phi$ by a real number: $\gamma \neq 0$, and that $V'_\alpha \neq V'_\beta$ for each $\alpha \in \Phi$. More precisely, $V'_\alpha \nsubseteq V'_\beta$ for each $\alpha \in \Phi$, so that $V'_\alpha \cap V'_\beta$ is a hyperplane in $V'_\alpha$ for every $\alpha \in \Phi$. It follows that the union of $V'_\alpha \cap V'_\beta$ over $\alpha \in \Phi$ is a proper subset of $V'_\alpha$, as in Section 18.12. Equivalently, this means that $V'_\alpha$ is not contained in the union of $V'_\alpha$, $\alpha \in \Phi$. Let $\tau$ be an element of $V'_\alpha$ that is not in $V'_\alpha$ for any $\alpha \in \Phi$. Thus $\tau \in V'$ satisfies

$$\tau(\gamma) = 0 \text{ and } \tau(\alpha) \neq 0 \text{ for every } \alpha \in \Phi.$$  

As in Section 19.14, there is an element $\sigma$ of the Weyl group of $\Phi$ such that $\tau(\sigma(\beta)) \geq 0$ for every $\beta \in \Delta$. This means that

$$\tau(\sigma(\beta)) > 0 \text{ for every } \beta \in \Delta,$$

because $\sigma(\beta) \in \Phi$, so that $\tau(\sigma(\beta)) \neq 0$, by (28.1.2). Note that $\sigma(\Delta)$ is a basis for $V$, because $\Delta$ is a basis for $V$. Thus we can express $\gamma$ as

$$\gamma = \sum_{\beta \in \Delta} c_\beta \sigma(\beta)$$

(28.1.4)
for some real numbers \( c_\beta, \beta \in \Delta \). It follows that

\[(28.1.5) \quad 0 = \tau(\gamma) = \sum_{\beta \in \Delta} c_\beta \tau(\sigma(\beta)).\]

Of course, \( c_\beta \neq 0 \) for some \( \beta \in \Delta \), because \( \gamma \neq 0 \). In fact, (28.1.5) implies that

\[(28.1.6) \quad c_\beta > 0 \text{ for some } \beta \in \Delta, \text{ and } c_\beta < 0 \text{ for some } \beta \in \Delta,
\]

because of (28.1.3). This corresponds to Exercise 10 on p54 of [13], and (g) on p54 of [23].

Suppose now that \( \gamma \) is also an element of the subgroup \( \Theta_\Phi \) of \( V \) generated by \( \Phi \), as a commutative group with respect to addition. This implies that

\[(28.1.7) \quad \sigma^{-1}(\gamma) = \sum_{\beta \in \Delta} c_\beta \beta
\]

is an element of \( \Theta_\Phi \) too, as in (27.14.1), so that \( c_\beta \in \mathbb{Z} \) for every \( \beta \in \Delta \). Let \( \nu_\gamma \in \mathcal{H}' \) be as in (27.14.2), and similarly for \( \nu_{\sigma^{-1}(\gamma)} \). Remember that \( L_\mu \subseteq L \) is defined for \( \mu \in \mathcal{H}' \) as in (27.11.1). It follows that

\[(28.1.8) \quad \dim L_{\nu_\gamma} = \dim L_{\nu_{\sigma^{-1}(\gamma)}},\]

as in (27.14.8).

As in Section 27.11, \( L_\mu = \{0\} \) unless \( \mu = 0 \), or \( \mu \) can be expressed as a sum of \( \nu_\beta \)'s, \( \beta \in \Delta \), or \( \mu \) can be expressed as a sum of \( -\nu_\beta \)'s, \( \beta \in \Delta \). This implies that

\[(28.1.9) \quad L_{\nu_{\sigma^{-1}(\gamma)}} = \{0\}\]

under the conditions described in the previous paragraphs, because of (28.1.6). This means that

\[(28.1.10) \quad L_{\nu_\gamma} = \{0\},\]

by (28.1.8). More precisely, this holds when \( \gamma \in \Theta_\Phi \) satisfies (28.1.1). This corresponds to part of Step (11) on p100 of [13], and of (h) on p54 of [23].

### 28.2 The dimension of \( L \)

Let us continue with the same notation and hypotheses as in the previous sections. If \( \alpha \in \Phi, \gamma \in \Theta_\Phi \), and

\[(28.2.1) \quad \gamma = t\alpha
\]

for some \( t \in \mathbb{R} \), then

\[(28.2.2) \quad t \in \mathbb{Z}.
\]

This is easy to see when \( \alpha \) is an element of a base for \( \Phi \), because \( \gamma \) can be expressed in a unique way as a linear combination of elements of the base, with coefficients in \( \mathbb{Z} \). Remember that \( \Phi \) is supposed to be reduced as a root system.
CHAPTER 28. SOME MORE LIE ALGEBRAS, 2

in $V$, as in Section 27.1. This implies that every element of $\Phi$ is an element of a base for $\Phi$, as in Section 19.14.

If (28.2.1) and (28.2.2) hold, then (28.1.10) holds when

$$|t| \geq 2,$$

as in (27.15.4). Combining this with the remarks in the previous section, we get that (28.1.10) holds for every $\gamma \in \Theta_\Delta$ such that $\gamma \neq 0$ and $\gamma \notin \Phi$.

Remember that $L_\mu = \{0\}$ unless $\mu \in H'$ satisfies $\mu = 0$, or $\mu$ is a sum of $\nu_\beta$'s, $\beta \in \Delta$, as in (27.13.7). It follows from this and the remarks in the preceding paragraph that $L_\mu = \{0\}$ unless $\mu = 0$, or $\mu = \nu_\alpha$ for some $\alpha \in \Phi$.

Remember that $X$ corresponds to the direct sum of the subspaces $L_\mu$, where $\mu \in H'$ can be expressed as a nontrivial sum of $\nu_\beta$'s, $\beta \in \Delta$, as a vector space over $k$, as in Section 27.11. Similarly, $Y$ corresponds to the direct sum of the subspaces $L_\mu$, where $\mu$ can be expressed as a nontrivial sum of $\nu_\beta$'s, $\beta \in \Delta$, as a vector space over $k$. Let $\Phi^+ = \Phi^+_\Delta$ be the set of $\alpha \in \Phi$ that can be expressed as a linear combination of $\beta \in \Delta$ with nonnegative coefficients, so that $\Phi = \Phi^+ \cup (-\Phi^+)$. Note that $\nu_{-\alpha} = -\nu_\alpha$ for every $\alpha \in \Phi$, by the definition (27.14.2) of $\nu_\alpha$. Thus $X$ corresponds to the direct sum of the subspaces $L_{\nu_\alpha}$, $\alpha \in \Phi^+$, as a vector space over $k$, and $Y$ corresponds to the direct sum of the subspaces $L_{\nu_\alpha}$, $\alpha \in \Phi^+$, as a vector space over $k$. These subspaces have dimension equal to 1, as in (27.15.3). This means that $X$ and $Y$ are finite-dimensional as vector spaces over $k$, with dimension equal to the number of elements of $\Phi^+$. Remember that $L$ corresponds to the direct sum of $X$, $H$, and $Y$ as a vector space over $k$, as in (27.10.1). We also have that the $h_\alpha$'s, $\alpha \in \Delta$, form a basis for $H$ as a vector space over $k$, as in Section 27.10. Thus the dimension of $H$ as a vector space over $k$ is equal to the number of elements of $\Delta$, which is the same as the dimension of $V$ as a vector space over $\mathbb{R}$. It follows that the dimension of $L$, as a vector space over $k$, is equal to the sum of the number of elements of $\Phi$ and the number of elements of $\Delta$. This corresponds to Step (12) on p100 of [13], and to (i) on p54 of [23].

28.3 Automorphisms and the Weyl group

We continue with the same notation and hypotheses as in the previous sections. If $\alpha \in \Phi$, then $\sigma_\alpha$ is the symmetry on $V$ with vector $\alpha$ that maps $\Phi$ onto itself, as usual. Suppose that $\alpha \in \Delta$, and let $\theta_\alpha$ be the automorphism of $L$ associated to $\alpha$ as in Section 27.12. Remember that $\theta_\alpha$ maps $H$ onto itself, as in (27.13.7). If $\mu \in H'$, then let $\zeta_{\mu,\alpha} \in H'$ be as in (27.13.11), so that $\zeta_{\mu,\alpha}(h) = \mu(\theta_\alpha(h))$ for every $h \in H$.

Suppose that $\gamma$ is an element of the subgroup $\Theta_\Phi$ of $V$, as a commutative group with respect to addition, generated by $\Phi$. Thus $\nu_\gamma \in H'$ can be defined as in (27.14.2), so that $\zeta_{\nu_\gamma,\alpha}$ can be defined as before. Remember that $\zeta_{\nu_\gamma,\alpha} = \nu_{\sigma_\alpha(\gamma)}$, as in (27.14.5).
28.3. AUTOMORPHISMS AND THE WEYL GROUP

Let $\alpha_1, \ldots, \alpha_r$ be a finite sequence of elements of $\Delta$, possibly with repetitions. Thus

\begin{equation}
\sigma = \sigma_{\alpha_1} \circ \cdots \circ \sigma_{\alpha_r}
\end{equation}

is an element of the Weyl group of $\Phi$, and

\begin{equation}
\theta = \theta_{\alpha_r} \circ \cdots \circ \theta_{\alpha_1}
\end{equation}

is a Lie algebra automorphism of $L$. If $\mu \in \mathcal{H}'$, then put

\begin{equation}
\zeta_{\mu, \theta}(h) = \mu(\theta(h))
\end{equation}

for every $h \in \mathcal{H}$, which defines a linear functional on $\mathcal{H}'$. Observe that

\begin{equation}
\zeta_{\nu, \theta} = \nu_{\sigma(\gamma)},
\end{equation}

because of the analogous statement for $\sigma_\alpha$ mentioned in the preceding paragraph.

If $\mu \in \mathcal{H}'$, then let $L_\mu$ be as in (27.11.1). Observe that $h \in \mathcal{H}$ and $u \in L$ satisfy

\begin{equation}
[h, u] = \zeta_{\mu, \theta}(h) u
\end{equation}

if and only if

\begin{equation}
[\theta(h), \theta(u)] = \mu(\theta(h)) \theta(u).
\end{equation}

This means that

\begin{equation}
\theta(L_{\zeta_{\mu, \theta}}) = L_\mu,
\end{equation}

which could also be obtained from (27.13.14). It follows that

\begin{equation}
\theta(L_{\nu_{\sigma(\gamma)}}) = L_{\nu_\gamma},
\end{equation}

by (28.3.4).

If $\beta \in \Delta$, then $L_{\nu_\beta}$ is spanned by $x_\beta$, as in Section 27.15. Similarly, one can check that $L_{-\nu_\beta}$ is spanned by $y_\beta$. This implies that $[L_{\nu_\beta}, L_{-\nu_\beta}]$ is spanned by $h_\beta$, because of (27.8.5). Remember that $x_\beta, y_\beta, h_\beta \neq 0$, as in Section 27.10. The linear subspace of $L$ spanned by $L_{\nu_\beta}$, $L_{-\nu_\beta}$, and $[L_{\nu_\beta}, L_{-\nu_\beta}]$ is a Lie subalgebra of $L$ isomorphic to $\mathfrak{sl}_2(k)$, because of (27.8.5), (27.8.6), (27.8.7) and the fact that $n(\beta, \beta) = 2$, as in Section 27.1.

If $\gamma \in \Phi$, then there is an element $\sigma$ of the Weyl group of $\Phi$ such that $\sigma(\gamma)$ is contained in $\Delta$, as in Section 19.14. Remember that $\sigma$ can be expressed as in (28.3.1), with $\alpha_1, \ldots, \alpha_r \in \Delta$, as in Section 19.14, so that we get a Lie algebra automorphism $\theta$ of $L$ as in (28.3.2). This means that (28.3.8) holds, and similarly

\begin{equation}
\theta(L_{-\nu_{\sigma(\gamma)}}) = L_{-\nu_\gamma}.
\end{equation}

It follows that

\begin{equation}
\theta([L_{\nu_{\sigma(\gamma)}}, L_{-\nu_{\sigma(\gamma)}}]) = \theta([L_{\nu_{\sigma(\gamma)}}, \theta(L_{-\nu_{\sigma(\gamma)}})]) = [L_{\nu_\gamma}, L_{-\nu_\gamma}].
\end{equation}
In particular,
\[ (28.3.11) \dim[L_{\nu_\alpha}, L_{-\nu_\alpha}] = 1, \]
because \([L_{\nu_{\beta}(\gamma)}, L_{-\nu_{\beta}(\gamma)}]\) is one-dimensional as a linear subspace of \(L\), as in the preceding paragraph.

We also get that the linear subspace of \(L\) spanned by \(L_{\nu_\alpha}\), \(L_{-\nu_\alpha}\), and \([L_{\nu_\alpha}, L_{-\nu_\alpha}]\) is a Lie subalgebra of \(L\) isomorphic to \(sl_2(k)\). More precisely, this subspace is the image under \(\theta\) of the linear subspace of \(L\) spanned by \(L_{\nu_{\beta}(\gamma)}\), \(L_{-\nu_{\beta}(\gamma)}\), and \([L_{\nu_{\beta}(\gamma)}, L_{-\nu_{\beta}(\gamma)}]\). The latter subspace of \(L\) is a Lie subalgebra isomorphic to \(sl_2(k)\), because \(\sigma(\gamma) \in \Delta\), as before. This corresponds to (j) on p55 of [23].

### 28.4 Semisimplicity of \(L\)

Let us continue with the same notation and hypotheses as in the previous sections. We would like to show that if \(C\) is an ideal in \(L\) that is commutative as a Lie algebra, then \(C = \{0\}\), as in Section 9.4.

Remember that \(L\) corresponds to the direct sum of the linear subspaces \(L_{\mu}\), \(\mu \in \mathcal{H}'\), as in Section 27.11. As a vector space over \(k\). More precisely, \(L\) corresponds to the direct sum of \(\mathcal{H} = L_0\) and the linear subspaces \(L_{\nu_\alpha}\), \(\alpha \in \Phi\), as in Section 28.2. If \(h \in \mathcal{H}\), then \(ad_h\) maps \(C\) into itself, because \(C\) is an ideal in \(L\). This implies that \(C\) corresponds to the direct sum of \(C \cap \mathcal{H}\) and the subspaces \(C \cap L_{\nu_\alpha}, \alpha \in \Phi\), as a vector space over \(k\).

If \(\alpha \in \Phi\), then the linear subspace of \(L\) spanned by
\[ (28.4.1) \quad L_{\nu_\alpha}, L_{-\nu_\alpha}, \text{ and } [L_{\nu_\alpha}, L_{-\nu_\alpha}] \]
is a Lie subalgebra of \(L\) isomorphic to \(sl_2(k)\) as a Lie algebra over \(k\), as in the previous section. The intersection of \(C\) with this Lie subalgebra is an ideal in this Lie subalgebra, and hence the intersection is trivial, because \(sl_2(k)\) is simple. In particular, \(C \cap L_{\nu_\alpha} = \{0\}\) for every \(\alpha \in \Phi\). It follows that \(C \subset \mathcal{H}\).

If \(\beta \in \Delta\), then \(x_\beta \in L_{\nu_\alpha}\), as in Section 27.15. Equivalently, \([h, x_\beta] = \nu_\beta(h)x_\beta\) for every \(h \in \mathcal{H}\). If \(h \in C\), then \([h, x_\beta] \in C\), because \(C\) is an ideal in \(L\). This implies that \([h, x_\beta] = 0\) for every \(h \in C\), by (27.10.3). Thus \(\nu_\beta(h) = 0\) for every \(h \in C\). Remember that the \(\nu_\beta\)'s, \(\beta \in \Delta\), form a basis for the dual of \(\mathcal{H}\), as a vector space over \(k\), as in Section 27.11. This implies that \(C = \{0\}\), as desired. This corresponds to Step (13) on p100 of [13], and (k) on p55 of [23].

Of course, \(\mathcal{H}\) is commutative as a Lie subalgebra of \(L\), and in particular \(\mathcal{H}\) is nilpotent as a Lie algebra. Let us check that \(\mathcal{H}\) is self-normalizing in \(L\), which means that \(\mathcal{H}\) is a Cartan subalgebra of \(L\). Suppose that \(u \in L\) is in the normalizer \(N_L(\mathcal{H})\) of \(\mathcal{H}\) in \(L\), so that \(ad_u\) maps \(\mathcal{H}\) into itself. As before, \(u\) can be expressed as the sum of an element of \(\mathcal{H}\) and elements of \(L_{\nu_\alpha}\), \(\alpha \in \Phi\). If the component of \(u\) in \(L_{\nu_\alpha}\) is not equal to \(0\) for some \(\alpha \in \Phi\), then the component of \([h_\alpha, u]\) in \(L_{\nu_\alpha}\) is nonzero, and \([h_\alpha, u]\) is not contained in \(\mathcal{H}\). Thus the component of \(u\) in \(L_{\nu_\alpha}\) is equal to \(0\) for every \(\alpha \in \Phi\) when \(u \in N_L(\mathcal{H})\). This means that \(N_L(\mathcal{H})\) is contained in \(\mathcal{H}\), as desired.
28.5. AN ISOMORPHISM THEOREM

Remember that $V$ is the vector space over the real numbers of positive finite dimension in which $\Phi$ is a root system. Consider the linear subspace of $V$, as a vector space over $Q$, generated by $\nu_\alpha$. It is easy to see that the mapping $\gamma \mapsto \nu_\alpha$ defined in Section 27.14 extends to an injective linear mapping from the subspace of $V$ just mentioned into $H'$, as a vector space over $Q$.

Remember that $L_\mu \neq \{0\}$ when $\mu \in H'$ and $\mu \neq \{0\}$ exactly when $\mu = \nu_\alpha$ for some $\alpha \in \Phi$, as in Section 28.2. Thus we get the usual relationship between $L$ and the set of $\mu \in H'$ such that $\mu \neq 0$ and $L_\mu \neq \{0\}$. This corresponds to Step (14) on p100 of [13], and (l) on p55 of [23].

28.5 An isomorphism theorem

Let $V_1, V_2$ be vector spaces over the real numbers of equal positive finite dimension, and let $\Phi_1, \Phi_2$ be reduced root systems in $V_1, V_2$, respectively. If $i = 1, 2$ and $\alpha, \beta \in \Phi_i$, then let

\[ n_i(\alpha, \beta) \in \mathbb{Z} \]

be as in (27.1.1), as usual. Let $\Delta_i$ be a base for $\Phi_i$ for $i = 1, 2$, and remember that

\[ -3 \leq n_i(\alpha, \beta) \leq 0 \]

for every $\alpha, \beta \in \Delta_i$ with $\alpha \neq \beta$, as in (27.1.2).

Let $k$ be a field of characteristic 0, and suppose that $(A_i, [\cdot, \cdot]_{A_i})$ is a finite-dimensional Lie algebra over $k$ for $i = 1, 2$ with the following properties. If $i = 1, 2$ and $\alpha, \beta \in \Delta_i$, then $x_{i, \alpha}, y_{i, \alpha},$ and $h_{i, \alpha}$ are elements of $A_i$. If $\alpha, \beta \in \Delta_i$, then

\[ [h_{i, \alpha}, h_{i, \beta}]_{A_i} = 0, \]

\[ [x_{i, \alpha}, y_{i, \beta}]_{A_i} = \delta_{\alpha, \beta} h_{i, \alpha}, \]

\[ [h_{i, \alpha}, x_{i, \beta}]_{A_i} = n_i(\beta, \alpha) x_{i, \beta}, \]

\[ [h_{i, \alpha}, y_{i, \beta}]_{A_i} = -n_i(\beta, \alpha) y_{i, \beta}, \]

where $\delta_{\alpha, \beta} \in k$ is equal to 1 when $\alpha = \beta$, and is equal to 0 when $\alpha \neq \beta$, as usual. If $\alpha \neq \beta$, then

\[ (\text{ad}_{A_i} x_{i, \alpha})^{n_i(\beta, \alpha)+1} (y_{i, \beta}) = (\text{ad}_{A_i} y_{i, \alpha})^{n_i(\beta, \alpha)+1} (x_{i, \beta}) = 0. \]

We also ask that $A_i$ be generated as a Lie algebra over $k$ by

\[ x_{i, \alpha}, y_{i, \alpha}, \text{ and } h_{i, \alpha}, \alpha \in \Delta_i, \]

and that the dimension of $A_i$ as a vector space over $k$ be equal to the sum of the number of elements of $\Phi_i$ and the number of elements of $\Delta_i$.

Of course, the number of elements of $\Delta_i$ is the same as the dimension of $V_i$. Suppose that $\phi$ is a one-to-one mapping from $\Delta_1$ onto $\Delta_2$ such that

\[ n_2(\phi(\alpha), \phi(\beta)) = n_1(\alpha, \beta) \]
for every $\alpha, \beta \in \Delta_1$. Thus $\phi$ has a unique extension to a one-to-one linear mapping from $V_1$ onto $V_2$ that maps $\Phi_1$ onto $\Phi_2$, as in Section 20.2. Conversely, if a one-to-one linear mapping from $V_1$ onto $V_2$ maps $\Phi_1$ onto $\Phi_2$ and $\Delta_1$ onto $\Delta_2$, then its restriction to $\Delta_1$ satisfies (28.5.9), as before.

Under these conditions, there is a unique Lie algebra isomorphism from $A_1$ onto $A_2$ with

$$x_{1,\alpha} \mapsto x_{2,\phi(\alpha)}, \quad y_{1,\alpha} \mapsto y_{2,\phi(\alpha)}, \quad h_{1,\alpha} \mapsto h_{2,\phi(\alpha)},$$

for every $\alpha \in \Delta_1$. This corresponds to Theorem 8’ on p50 of [23], and to part (b) of the theorem on p101 of [13]. This is also related to Theorem 7 on p49 of [23], whose proof is discussed on p55 of [23]. Another approach to isomorphisms like these is discussed in Section 14.2 of [13].

Note that Serre’s theorem shows that there are Lie algebras with these properties, as in Sections 27.8 and 28.2. This permits us to reduce to the case where $A_1$ is as in Serre’s theorem. It follows from the proof of Serre’s theorem that there is a unique Lie algebra homomorphism from $A_1$ into $A_2$ that satisfies (28.5.10). This homomorphism is surjective, because $A_2$ is generated by (28.5.8) with $i = 2$, as a Lie algebra over $k$. This homomorphism is injective as well, because $A_1$ and $A_2$ have the same dimension, as vector spaces over $k$.

### 28.6 Some automorphisms

Let $V_1$ be a vector space over the real numbers of positive finite dimension, and let $\Phi_1$ be a reduced root system in $V_1$. Also let $n_1(\alpha, \beta)$ be as before for $\alpha, \beta \in \Phi_1$, and let $\Delta_1$ be a base for $\Phi_1$. Suppose that $(A_1,\{.,.\}_{A_1})$ is a finite-dimensional Lie algebra over a field $k$ of characteristic 0 with the properties mentioned in the previous section. In particular, $A_1$ should be generated as a Lie algebra over $k$ by (28.5.8), with $i = 1$.

In this section, we take $V_2 = V_1$, $\Phi_2 = \Phi_1$, and $A_2 = A_1$. Thus $n_2(.,.)$ is the same as $n_1(.,.)$. Let $\Delta_2$ be another base for $\Phi_1$, and suppose that we have generators for $A_1 = A_2$ as in (28.5.8), with $i = 2$, and which satisfy the same conditions as before.

Let $\phi$ be a one-to-one mapping from $\Delta_1$ onto $\Delta_2$ such that

$$n_1(\phi(\alpha), \phi(\beta)) = n_1(\alpha, \beta)$$

for every $\alpha, \beta \in \Delta_1$, which is the same as (28.5.9) in this case. As before, $\phi$ has a unique extension to an automorphism of $\Phi_1$. Conversely, if $T$ is any automorphism of $\Phi_1$, then $\Delta_2 = T(\Delta_1)$ is a base for $\Phi_1$, and the restriction of $T$ to $\Delta_1$ satisfies (28.6.1).

The isomorphism theorem described in the previous section implies that there is a unique Lie algebra automorphism of $A_1$ that satisfies (28.5.10). This is related to the remarks at the beginning of Section 14.3 on p76 of [13].

Remember that the mapping $T$ from $V_1$ onto itself defined by

$$T(v_1) = -v_1$$

for every $v_1$. Thus $\phi$ has a unique extension to a one-to-one linear mapping from $V_1$ onto $V_2$ that maps $\Phi_1$ onto $\Phi_2$, as in Section 20.2. Conversely, if a one-to-one linear mapping from $V_1$ onto $V_2$ maps $\Phi_1$ onto $\Phi_2$ and $\Delta_1$ onto $\Delta_2$, then its restriction to $\Delta_1$ satisfies (28.5.9), as before.
28.6. SOME AUTOMORPHISMS

defines an automorphism of \( \Phi_1 \). In particular,

\[
    n_1(-\alpha, -\beta) = n_1(\alpha, \beta)
\]

for every \( \alpha, \beta \in \Phi_1 \). If \( \alpha \in \Delta_1 \), then \( -\alpha \in T(\Delta_1) \), and we put

\[
    x_{2,-\alpha} = -y_{1,\alpha}, \quad y_{2,-\alpha} = -x_{1,\alpha}, \quad h_{2,-\alpha} = -h_{1,\alpha}.
\]

One can check that \( x_{2,-\alpha}, y_{2,-\alpha}, \) and \( h_{2,-\alpha}, \alpha \in \Delta_1 \), satisfy the requirements in the previous section for \( A_1 = A_2 \), using (28.6.3) and the analogous properties of \( x_{1,\alpha}, y_{1,\alpha}, \) and \( h_{1,\alpha} \).

It follows that there is a unique Lie algebra automorphism \( \sigma \) of \( A_1 \) with

\[
    x_{1,\alpha} \mapsto -y_{1,\alpha}, \quad y_{1,\alpha} \mapsto -x_{1,\alpha}, \quad h_{1,\alpha} \mapsto -h_{1,\alpha}
\]

for every \( \alpha \in \Delta_1 \). Observe that \( \sigma \circ \sigma \) is the identity mapping on \( A_1 \), because it sends the generators \( x_{1,\alpha}, y_{1,\alpha}, \) and \( h_{1,\alpha} \) to themselves for every \( \alpha \in \Delta_1 \). This corresponds to the corollary to Theorem 7 on p49 of [23], and to the remark after Theorem 8' on p50 of [23]. This also corresponds to the proposition on p77 of [13].

Of course, the identity mapping on \( V_1 \) is an automorphism of \( \Phi_1 \). Let \( t_\alpha \) be a nonzero element of \( k \) for every \( \alpha \in \Delta_1 \), and put \( \Delta_2 = \Delta_1 \) and

\[
    x_{2,\alpha} = t_\alpha x_{1,\alpha}, \quad y_{2,\alpha} = (1/t_\alpha) y_{1,\alpha}, \quad h_{2,\alpha} = h_{1,\alpha}
\]

for every \( \alpha \in \Delta_1 \). It is easy to see that \( x_{2,\alpha}, y_{2,\alpha}, \) and \( h_{2,\alpha}, \alpha \in \Delta_1 \), satisfy the requirements in the previous section for \( A_1 = A_2 \), using the analogous properties of \( x_{1,\alpha}, y_{1,\alpha}, \) and \( h_{1,\alpha} \). This implies that there is a unique Lie algebra automorphism of \( A_1 \) with

\[
    x_{1,\alpha} \mapsto t_\alpha x_{1,\alpha}, \quad y_{1,\alpha} \mapsto (1/t_\alpha) y_{1,\alpha}, \quad h_{1,\alpha} \mapsto h_{1,\alpha}
\]

for every \( \alpha \in \Delta_1 \). This corresponds to the diagonal automorphisms mentioned on p87 of [13].

Let \( T \) be an automorphism of \( \Phi_1 \) such that

\[
    T(\Delta_1) = \Delta_1.
\]

Put \( \Delta_2 = \Delta_1 \), and let \( \phi \) be the restriction of \( T \) to \( \Delta_1 \). Equivalently, one can take \( \phi \) to be any one-to-one mapping from \( \Delta_1 \) onto itself that satisfies (28.6.1).

Of course,

\[
    x_{2,\alpha} = x_{1,\alpha}, \quad y_{2,\alpha} = y_{1,\alpha}, \quad h_{2,\alpha} = h_{1,\alpha}
\]

automatically satisfy the requirements in the previous section for \( A_1 = A_2 \), by hypothesis. Thus there is a unique Lie algebra automorphism of \( A_1 \) with

\[
    x_{1,\alpha} \mapsto x_{1,\phi(\alpha)}, \quad y_{1,\alpha} \mapsto y_{1,\phi(\alpha)}, \quad h_{1,\alpha} \mapsto h_{1,\phi(\alpha)}
\]

for every \( \alpha \in \Delta_1 \). More precisely, the collection of automorphisms of \( \Phi_1 \) that satisfy (28.6.8) is a group with respect to composition of mappings, and we get an injective homomorphism from this group into the automorphism group of \( A_1 \). This corresponds to Exercise 6 on p77 of [13].
28.7 Roots and diagonalizability

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $\Phi$ be a reduced root system in $V$. As usual, if $\alpha \in \Phi$, then we let $\sigma_\alpha$ be the symmetry on $V$ with vector $\alpha$ that maps $\Phi$ onto itself. Thus $\sigma_\alpha$ can be expressed as $\sigma_\alpha(v) = v - \lambda_\alpha(v)\alpha$, where $\lambda_\alpha$ is a linear functional on $V$ with $\lambda_\alpha(\alpha) = 2$. Remember that $\lambda_\alpha(\beta) \in \mathbb{Z}$ for every $\beta \in \Phi$, by the definition of a root system.

Let $\Theta$ be the subgroup of $V$, as a commutative group with respect to addition, generated by $\Phi$. We would like to consider a Lie algebra $(A, [\cdot, \cdot])$ over a field $k$ with properties like those mentioned in Sections 22.11, 23.3, 23.4, and 23.5. In particular, these conditions will hold in the situation described in Section 22.1, as before.

If $\alpha \in \Phi \cup \{0\}$, then we ask that $A_\alpha$ be a linear subspace of $A$, and that $A$ correspond to the direct sum of $A_\alpha$, $\alpha \in \Phi \cup \{0\}$, as a vector space over $k$. More precisely, if $\alpha \in \Phi$, then we ask that $A_\alpha$ have dimension one. As before, it is convenient to put $A_0 = \{0\}$ when $\alpha = 0$.

Suppose that $x_{A,\alpha}, y_{A,\alpha}$ are nonzero elements of $A_\alpha$, $A_{-\alpha}$, respectively, for each $\alpha \in \Delta$, and put

\[ h_{A,\alpha} = [x_{A,\alpha}, y_{A,\alpha}], \]

which is an element of $A_0$. We ask that $\phi_\beta(h_{A,\alpha}) = \lambda_\alpha(\beta) \cdot 1$ in $k$ for every $\alpha \in \Delta$ and $\beta \in \Phi$, and that the $h_{A,\alpha}$'s, $\alpha \in \Delta$, form a basis for $A_0$.

Under these conditions, $A$ is generated by $x_{A,\alpha}, y_{A,\alpha}, h_{A,\alpha}, \alpha \in \Delta$, as a Lie algebra over $k$, as in Section 23.4. These generators satisfy the same relations as in Section 28.5, as discussed in Section 23.5. We also have that the dimension of $A$ as a vector space over $k$ is equal to the sum of the numbers of elements of $\Phi$ and $\Delta$.

If $k$ has characteristic 0, then all of the conditions mentioned in Section 28.5 are satisfied. It follows that there is a unique Lie algebra isomorphism from the Lie algebra $L$ obtained from Serre’s theorem in Section 27.8 onto $A$, with

\[ x_\alpha \mapsto x_{A,\alpha}, \quad y_\alpha \mapsto y_{A,\alpha}, \quad h_\alpha \mapsto h_{A,\alpha} \]

for every $\alpha \in \Delta$. This corresponds to Theorem 7 on p49 of [23]. This is also related to Theorems 8 and 8’ on p50 of [23], and part (b) of the theorem on p101 of [13]. Note that the remarks in the previous section can be used as well in this case.
28.8 Quotients of $\tilde{L}$

Let $V$ be a vector space over the real numbers of positive finite dimension again, and let $\Phi$ be a reduced root system in $V$. Also let $\sigma_\alpha, \lambda_\alpha$ be as before for every $\alpha \in \Phi$, and put $\sigma(\alpha, \beta) = \lambda_\beta(\alpha)$ for every $\alpha, \beta \in \Phi$, as in (27.1.1). Let $\Delta$ be a base for $\Phi$ as well, and let $k$ be a field of characteristic 0. Under these conditions, the Lie algebra $\mathfrak{e}L = \mathfrak{e}L(k)$ can be defined as in Section 27.1.

Let $(B, [\cdot, \cdot]_B)$ be a Lie algebra over $k$ with the following properties. If $\alpha \in \Delta$, then there are elements $x_{B,\alpha}, y_{B,\alpha}, h_{B,\alpha}$ that satisfy the Weyl relations. This means that for every $\alpha, \beta \in \Delta$, we have that

\begin{align}
&[h_{B,\alpha}, b_{B,\beta}]_B = 0, \\
&[x_{B,\alpha}, y_{B,\beta}]_B = \delta_{\alpha,\beta} h_{B,\alpha}, \\
&[h_{B,\alpha}, x_{B,\beta}]_B = n(\beta, \alpha) x_{B,\beta}, \\
&[h_{B,\alpha}, y_{B,\beta}]_B = -n(\beta, \alpha) y_{B,\beta}.
\end{align}

Let us also ask that $B$ be generated by $x_{B,\alpha}, y_{B,\alpha}, h_{B,\alpha}, \alpha \in \Delta$, as a Lie algebra over $k$. Of course, this can be arranged by replacing $B$ with the subalgebra generated by these elements.

Remember that $\tilde{L}$ is generated by elements $\tilde{x}_\alpha, \tilde{y}_\alpha, \tilde{h}_\alpha, \alpha \in \Delta$, that satisfy the Weyl relations. More precisely, $\tilde{L}$ was defined to be the quotient of a suitable free Lie algebra by the ideal corresponding to the Weyl relations. It follows that there is a unique Lie algebra homomorphism from $\tilde{L}$ onto $B$ with

$$
\tilde{x}_\alpha \mapsto x_{B,\alpha}, \quad \tilde{y}_\alpha \mapsto y_{B,\alpha}, \quad \tilde{h}_\alpha \mapsto h_{B,\alpha}
$$

for every $\alpha \in \Delta$.

Let $\alpha, \beta \in \Delta$ and a nonnegative integer $l$ be given. Observe that

\begin{align}
&\text{ad}_{\tilde{L},\tilde{h}_\alpha} \left( (\text{ad}_{\tilde{L},\tilde{x}_\alpha})^l(\tilde{x}_\beta) \right) = (l \cdot n(\alpha, \alpha) + n(\beta, \alpha)) (\text{ad}_{\tilde{L},\tilde{x}_\alpha})^l(\tilde{x}_\beta) \\
&\text{ad}_{\tilde{L},\tilde{y}_\alpha} \left( (\text{ad}_{\tilde{L},\tilde{x}_\alpha})^l(\tilde{y}_\beta) \right) = -(l \cdot n(\alpha, \alpha) + n(\beta, \alpha)) (\text{ad}_{\tilde{L},\tilde{x}_\alpha})^l(\tilde{y}_\beta),
\end{align}

by (27.4.2) and (27.4.3), respectively. This implies that

\begin{align}
&\text{ad}_{B,h_{B,\alpha}} ((\text{ad}_{B,x_{B,\alpha}})^l(x_{B,\beta})) \\
&\quad = (l \cdot n(\alpha, \alpha) + n(\beta, \alpha)) (\text{ad}_{B,x_{B,\alpha}})^l(x_{B,\beta})
\end{align}

and

\begin{align}
&\text{ad}_{B,h_{B,\alpha}} ((\text{ad}_{B,y_{B,\alpha}})^l(y_{B,\beta})) \\
&\quad = -(l \cdot n(\alpha, \alpha) + n(\beta, \alpha)) (\text{ad}_{B,y_{B,\alpha}})^l(y_{B,\beta}).
\end{align}

Of course, these identities could also be obtained more directly, using the same type of argument.
Remember that \( n(\alpha, \alpha) = \lambda_\alpha(\alpha) = 2 \). If, for each \( l \geq 0 \),
\[
(\text{ad}_{B,x,B})^l(x_{B,\beta}) \neq 0,
\]
then it follows that \( B \) has infinite dimension as a vector space over \( k \), because these elements of \( B \) are eigenvectors of \( \text{ad}_{B,h,B} \) with distinct eigenvalues. Equivalently, if \( B \) has finite dimension as a vector space over \( k \), then
\[
(\text{ad}_{B,x,B})^l(x_{B,\beta}) = 0
\]
for some \( l \geq 0 \). Similarly,
\[
(\text{ad}_{B,y,B})^l(y_{B,\beta}) = 0
\]
for some \( l \geq 0 \) in this case. This corresponds to the fact that (ii) implies (iii) in the exercise on p55 of [23], which is related to Exercise 3 on p101 of [13].

### 28.9 Additional identities in \( B \)

Let us continue with the same notation and hypotheses as in the previous section. Let \( \alpha, \beta \in \Delta \) be given again, and suppose that \( \alpha \neq \beta \). Remember that
\[
(\text{ad}_{L,x,a}^{-1})^j(\bar{y}_\gamma) = j (-n(\beta, \alpha) - j + 1) (\text{ad}_{L,y,a}^{-1})^{j-1}(\bar{y}_\gamma)
\]
for every positive integer \( j \), as in (27.7.9). Similarly,
\[
(\text{ad}_{L,y,a}^{-1})^j(\bar{x}_\beta) = j (-n(\beta, \alpha) - j + 1) (\text{ad}_{L,x,a}^{-1})^{j-1}(\bar{x}_\beta)
\]
for every positive integer \( j \). This can be obtained from (28.9.1) using the automorphism \( \bar{\sigma} \) of \( \tilde{L} \) discussed in Section 27.6, which sends \( \bar{x}_\gamma, \bar{y}_\gamma \) to \( -\bar{y}_\gamma, -\bar{x}_\gamma \), respectively, for each \( \gamma \in \Delta \).

It follows that
\[
\text{ad}_{B,x,B,a}((\text{ad}_{B,y,B,a})^j(y_{B,\beta})) = j (-n(\beta, \alpha) - j + 1) (\text{ad}_{B,y,B,a})^{j-1}(y_{B,\beta})
\]
for every positive integer \( j \), using the Lie algebra homomorphism from \( \tilde{L} \) onto \( B \) mentioned in the previous section. Similarly,
\[
\text{ad}_{B,y,B,a}((\text{ad}_{B,x,B,a})^j(x_{B,\beta})) = j (-n(\beta, \alpha) - j + 1) (\text{ad}_{B,x,B,a})^{j-1}(x_{B,\beta})
\]
for every positive integer \( j \).

Remember that \(-3 \leq n(\beta, \alpha) \leq 0\), as in (27.1.2). If (28.8.11) holds for some \( l \geq 0 \), then
\[
(\text{ad}_{B,x,B,a})^{-n(\beta, \alpha)+1}(x_{B,\beta}) = 0.
\]
More precisely, this is trivial when
\[(28.9.6) \quad l \leq -n(\beta, \alpha) + 1,\]
and otherwise one can use (28.9.4). Similarly, if (28.8.12) holds for some \(l \geq 0\), then
\[(28.9.7) \quad (\text{ad}_{B,y_B,\alpha})^{-n(\beta, \alpha)+1}(y_{B,\beta}) = 0,\]
because of (28.9.3). This corresponds to the fact that (iii) implies (i) in the exercise on p55 of [23]. In particular, (28.9.5) and (28.9.7) hold when \(B\) has finite dimension as a vector space over \(k\), by the remarks in the previous section. This corresponds to Exercise 3 on p101 of [13].

Let \(L = L(k)\) be the Lie algebra defined as in Serre’s theorem, in Section 27.8. Remember that \(L\) is a quotient of \(\tilde{L}\), and that the natural quotient mapping sends \(\tilde{x}_\gamma, \tilde{y}_\gamma,\) and \(\tilde{h}_\gamma\) to \(x_\gamma, y_\gamma,\) and \(h_\gamma\), respectively, for each \(\gamma \in \Delta\). If (28.9.5) and (28.9.7) hold for every \(\alpha, \beta \in \Delta\) with \(\alpha \neq \beta\), then there is a unique Lie algebra homomorphism from \(L\) onto \(B\) with
\[(28.10.1) \quad x_\gamma \mapsto x_{B,\gamma}, \quad y_\gamma \mapsto y_{B,\gamma}, \quad h_\gamma \mapsto h_{B,\gamma}\]
for every \(\gamma \in \Delta\). More precisely, the Lie algebra homomorphism from \(\tilde{L}\) onto \(B\) mentioned in the previous section is the same as the composition of the natural quotient mapping from \(\tilde{L}\) onto \(L\) with this homomorphism from \(L\) onto \(B\). This shows that any finite-dimensional quotient of \(\tilde{L}\) is in fact a quotient of \(L\).

28.10 Eigenspaces in \(L\)

Let \(V\) be a vector space over the real numbers of positive finite dimension, and let \(\Phi\) be a reduced root system in \(V\) again. If \(\alpha, \beta \in \Phi\), then we let \(n(\alpha, \beta) \in \mathbb{Z}\) be as in (27.1.1), and we let \(\Delta\) be a base for \(\Phi\). Let \(k\) be a field of characteristic 0, and let \(L = L(k)\) be the Lie algebra over \(k\) defined as in Serre’s theorem, in Section 27.8. Remember that \(L\) is generated by the elements \(x_\alpha, y_\alpha,\) and \(h_\alpha\) with \(\alpha \in \Delta\).

Let \(H\) be the linear subspace of \(L\) spanned by the \(h_\alpha\)’s, \(\alpha \in \Delta\), as in Section 27.10. Also let \(H'\) be the dual of \(H\), as a vector space over \(k\). If \(\mu \in H'\), then let \(L_\mu\) be the set of \(u \in L\) such that
\[(28.10.1) \quad [h, u] = \mu(h) u\]
for every \(h \in H\), as in (27.11.1). If \(\mu_1, \mu_2 \in H'\), then it is easy to see that
\[(28.10.2) \quad [L_{\mu_1}, L_{\mu_2}] \subseteq L_{\mu_1 + \mu_2},\]
using the Jacobi identity. Remember that \(H = L_0\), as in (27.11.4).

Let \(\Theta_\Phi\) be the subgroup of \(V\), as a commutative group with respect to addition, generated by \(\Phi\), which is the same as the subgroup generated by \(\Delta\).
If $\gamma \in \Theta_\Phi$, then let $\nu_\gamma$ be the linear functional defined on $\mathcal{H}$ as in (27.14.2). If $\gamma \in \Phi$, then

\begin{equation}
\nu_\gamma(h_\beta) = n(\gamma, \beta)
\end{equation}

for every $\beta \in \Delta$, because of (27.1.1). Remember that $\gamma \mapsto \nu_\gamma$ is a homomorphism from $\Theta_\Phi$ into $\mathcal{H}'$, as commutative groups with respect to addition.

If $\gamma \in \Phi$, then $L_{\nu_\gamma}$ has dimension one as a vector space over $k$, as in Section 27.15. Let $\gamma \in \Phi$ be given, and let $\sigma$ be an element of the Weyl group of $\Phi$ such that $\sigma(\gamma) \in \Delta$, as in Section 19.14. This leads to an automorphism $\theta$ of $L$, as in Section 28.3. By construction, $\theta$ maps $L_{\nu_{\sigma(\gamma)}}$, $L_{-\nu_{\sigma(\gamma)}}$ onto $L_{\nu_\gamma}$, $L_{-\nu_\gamma}$, respectively, and thus $[L_{\nu_{\sigma(\gamma)}}, L_{-\nu_{\sigma(\gamma)}}]$ onto $[L_{\nu_\gamma}, L_{-\nu_\gamma}]$, as before.

Remember that $[L_{\nu_{\sigma(\gamma)}}, L_{-\nu_{\sigma(\gamma)}}]$ is spanned by $h_{\sigma(\gamma)}$, because $\sigma(\gamma) \in \Delta$. Put

\begin{equation}
h_\gamma = \theta(h_{\sigma(\gamma)}),
\end{equation}

which is an element of $[L_{\nu_\gamma}, L_{-\nu_\gamma}]$. In fact, $[L_{\nu_\gamma}, L_{-\nu_\gamma}]$ is spanned by $h_\gamma$, because $[L_{\nu_{\sigma(\gamma)}}, L_{-\nu_{\sigma(\gamma)}}]$ is spanned by $h_{\sigma(\gamma)}$.

If $\alpha \in \Theta_\Phi$, then

\begin{equation}
\nu_\alpha(h_\gamma) = \nu_\alpha(\theta(h_{\sigma(\gamma)})) = \nu_{\sigma(\alpha)}(h_{\sigma(\gamma)}),
\end{equation}

using (28.3.4) in the second step. If $\alpha \in \Phi$, then $\sigma(\alpha) \in \Phi$, then we get that

\begin{equation}
\nu_\alpha(h_\gamma) = n(\sigma(\alpha), \sigma(\gamma)),
\end{equation}

by (28.10.3), and because $\sigma(\gamma) \in \Delta$. This implies that

\begin{equation}
\nu_\alpha(h_\gamma) = n(\alpha, \gamma),
\end{equation}

because $\sigma$ is an automorphism of $\Phi$, as in Section 20.2.

In particular,

\begin{equation}
\nu_\gamma(h_\gamma) = 2,
\end{equation}

by the definition of $n(\gamma, \gamma)$, as in (27.1.1). It follows that $h_\gamma$ does not depend on the choice of $\sigma$. More precisely, this also uses the fact that $[L_{\nu_\gamma}, L_{-\nu_\gamma}]$ is spanned by $h_\gamma$, as before.

If $\gamma_1, \gamma_2 \in \Phi$ and $\gamma_1 + \gamma_2 \in \Phi$, then

\begin{equation}
[L_{\nu_{\gamma_1}}, L_{\nu_{\gamma_2}}] = L_{\nu_{\gamma_1 + \gamma_2}}.
\end{equation}

Remember that $L$ is semisimple, as in Section 28.4. This permits us to use many of the remarks in Chapter 17, although a number of these properties have already been established. In particular, (28.10.9) was discussed in Section 17.9. Remember that if $\gamma \in \Phi$, then the linear subspace of $L$ spanned by $L_{\nu_\gamma}$, $L_{-\nu_\gamma}$, and $[L_{\nu_\gamma}, L_{-\nu_\gamma}]$ is a Lie subalgebra of $L$ that is isomorphic to $sl_2(k)$, as in Section 28.3.

Let $\alpha \in \Delta$ be given, and remember that $x_\alpha \in L_{\nu_\alpha}$, $y_\alpha \in L_{-\nu_\alpha}$, by construction. If $\gamma \in \Phi$ and $\alpha + \gamma \in \Phi$, then the restriction of $\text{ad}_{x_\alpha}$ to $L_{\nu_\alpha}$ is a one-to-one mapping onto $L_{\nu_{\alpha + \gamma}}$. Similarly, the restriction of $\text{ad}_{y_\alpha}$ to $L_{-\nu_\alpha}$ is a one-to-one
28.11. SUBSETS OF $\Delta$

mapping onto $L_{-\alpha, -\gamma}$. This is the same as (28.10.9) with $\gamma_1 = \alpha$, $\gamma_2 = \gamma$ and $\gamma_1 = -\alpha$, $\gamma_2 = -\gamma$, respectively, as in Section 17.9. Note that the fact that the linear subspace of $L$ spanned by $x_\alpha$, $y_\alpha$, and $h_\alpha$ is a Lie subalgebra of $L$ isomorphic to $sl_2(k)$ can be verified more directly, as in Section 28.3.

If $u_1, \ldots, u_r \in L$ for some $r \geq 2$, then put

$$[u_1, \ldots, u_r] = [u_1, [u_2, \ldots, [u_{r-1}, u_r]\ldots]],$$

as before. Let $\alpha_1, \ldots, \alpha_r$ be a finite sequence of elements of $\Delta$, possible with repetitions, such that

$$\sum_{j=1}^l \alpha_j \in \Phi$$

for every $l = 1, \ldots, r$. If $r \geq 2$, then we get that

$$[x_{\alpha_1}, \ldots, x_{\alpha_r}] \neq 0$$

and

$$[y_{\alpha_1}, \ldots, y_{\alpha_r}] \neq 0,$$

as in the preceding paragraph.

28.11 Subsets of $\Delta$

Let us continue with the same notation and hypotheses as in the previous section. Let $\Delta_0$ be a nonempty subset of $\Delta$, and let $V_0$ be the linear subspace of $V$ spanned by $\Delta_0$. Remember that

$$\Phi_0 = \Phi \cap V_0$$

is a root system in $V$, as in Section 20.6. If $\alpha, \beta \in \Phi_0$, then the analogue $n_0(\alpha, \beta)$ of (27.1.1) for $\Phi_0$ is the same as $n(\alpha, \beta)$, as in Section 20.15. Note that $\Delta_0$ is a base for $\Phi_0$, as before.

Let $L^0 = L^0(k)$ be the Lie algebra over $k$ defined as in Serre’s theorem in Section 27.8, using $\Phi_0$ and $\Delta_0$. More precisely, if $\alpha \in \Delta_0$, then let $x_{\alpha_0}, y_{\alpha_0},$ and $h_{0, \alpha}$ be the corresponding generators in $L^0$, as before. These generators satisfy the same relations in $L^0$ as their analogues in $L$, because $n_0(\alpha, \beta) = n(\alpha, \beta)$ for every $\alpha, \beta \in \Delta_0$.

The construction used in Serre’s theorem implies that there is a unique Lie algebra homomorphism from $L^0$ into $L$ with

$$x_{0, \alpha} \mapsto x_\alpha, \quad y_{0, \alpha} \mapsto y_\alpha, \quad h_{0, \alpha} \mapsto h_\alpha,$$

for every $\alpha \in \Delta_0$. This homomorphism maps $L^0$ onto the Lie subalgebra of $L$ generated by $x_\alpha, y_\alpha, h_\alpha$, with $\alpha \in \Delta_0$. We would like to check that this homomorphism is injective. This corresponds to Exercise 8 on p101 of [13].

It suffices to show that the dimension of the Lie subalgebra of $L$ generated by $x_\alpha, y_\alpha, h_\alpha$, $\alpha \in \Delta_0$, is at least the dimension of $L^0$. To do this, it is enough
to verify that if $\gamma \in \Phi_0$, then $L_{\nu}$ is contained in this subalgebra of $L$. This is clear when $\gamma \in \Delta_0$ or $-\gamma \in \Delta_0$.

Let $\gamma \in \Phi_0$ be a positive root with respect to $\Delta_0$, so that $\gamma$ can be expressed as a linear combination of elements of $\Delta_0$ with nonnegative coefficients. Remember that $\gamma$ can be expressed as $\sum_{j=1}^{r} \alpha_j$, where $\alpha_j \in \Delta_0$ for every $j = 1, \ldots, r$, and

$$\sum_{j=1}^{l} \alpha_j \in \Phi_0$$

for every $l = 1, \ldots, r$, as in Section 19.12. Suppose that $\gamma \notin \Delta_0$, so that $r \geq 2$.

Observe that

$$[x_{\alpha_1}, \ldots, x_{\alpha_r}]$$

and

$$[y_{\alpha_1}, \ldots, y_{\alpha_r}]$$

are elements of this subalgebra of $L$. These are also elements of $L_{\nu}$ and $L_{-\nu}$, respectively.

Remember that (28.11.4) and (28.11.5) are nonzero under these conditions, as in the previous section. It follows that $L_{\nu}$ and $L_{-\nu}$ are spanned by (28.11.4) and (28.11.5), respectively, because $L_{\nu}$ and $L_{-\nu}$ are one-dimensional linear subspaces of $L$, as in Section 27.15. This implies that $L_{\nu}$ and $L_{-\nu}$ are contained in this subalgebra of $L$, as desired.
Chapter 29

Chevalley’s normalization

29.1 Linearly independent roots

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $\Phi$ be a reduced root system in $V$. As usual, if $\alpha \in \Phi$, then we let $\sigma_\alpha$ be the symmetry on $V$ with vector $\alpha$ that maps $\Phi$ onto itself. Let $\langle \cdot, \cdot \rangle$ be an inner product on $V$ that is invariant under the Weyl group of $\Phi$, so that $\sigma_\alpha$ is the reflection on $V$ with respect to $\langle \cdot, \cdot \rangle$ associated to $\alpha$ for every $\alpha \in \Phi$. If $\alpha, \beta \in \Phi$, then

$$n(\beta, \alpha) = 2\langle \beta, \alpha \rangle \langle \alpha, \alpha \rangle^{-1} \in \mathbb{Z},$$

as before.

Suppose that $\alpha, \beta \in \Phi$ are linearly independent in $V$, which is to say that they are not proportional in $V$. Let $r, q$ be the largest nonnegative integers such that

$$\beta - r\alpha, \beta + q\alpha \in \Phi,$$

as in Section 20.5. Remember that

$$2\langle \beta, \alpha \rangle \langle \alpha, \alpha \rangle^{-1} = r - q,$$

as in (20.5.10). This corresponds to part (a) of the proposition on p145 of [13], and was also mentioned on p45 of [13].

Let $V_1$ be the linear subspace of $V$ spanned by $\alpha$ and $\beta$, and let $\Phi_1$ be the set of elements of $\Phi$ that can be expressed as a linear combination of $\alpha$ and $\beta$ with integer coefficients. Note that if the sum of two elements of $\Phi_1$ is an element of $\Phi$, then the sum is also an element of $\Phi_1$. It follows that $\Phi_1$ is a root system in $V_1$, as in Section 20.6. More precisely, $\Phi_1$ is reduced, because $\Phi$ is reduced, by hypothesis. This corresponds to Exercise 7 on p46 of [13], as before.

Under these conditions, there are at most two possible values for the norms of elements of $\Phi_1$ with respect to $\langle \cdot, \cdot \rangle$, as in part (b) of the proposition on p145 of [13], and its proof. More precisely, this follows from the remarks in Section 20.7 when $\Phi_1$ is irreducible in $V_1$. Otherwise, if $\Phi_1$ is reducible, then it consists
of only \( \pm \alpha, \pm \beta \), because \( \Phi_1 \) is reduced. In this case, it is clear that there are at most two possible values for the norms of elements of \( \Phi_1 \).

If \( \alpha + \beta \in \Phi \), then part (c) of the proposition on p145 of [13] states that

\[
(29.1.4) \quad r + 1 = q (\alpha + \beta, \alpha + \beta) (\beta, \beta)^{-1}.
\]

To see this, we first use (29.1.3) to get that

\[
(29.1.5) \quad r + 1 - q (\alpha + \beta, \alpha + \beta) (\beta, \beta)^{-1} = q + 2 (\beta, \alpha) (\alpha, \alpha)^{-1} + 1 - q (\alpha + \beta, \alpha + \beta, \beta, \beta)^{-1} = 2 (\beta, \alpha) (\alpha, \alpha)^{-1} + 1 - q (\alpha, \alpha) (\beta, \beta)^{-1} - 2 q (\alpha, \beta) (\beta, \beta)^{-1}.
\]

If we put

\[
(29.1.6) \quad A = 2 (\beta, \alpha) (\alpha, \alpha)^{-1} + 1, \\
(29.1.7) \quad B = 1 - q (\alpha, \alpha) (\beta, \beta)^{-1},
\]

then the right side of (29.1.5) is the same as \( AB \). Thus it suffices to show that \( A = 0 \) or \( B = 0 \).

Remember that

\[
(29.1.8) \quad n(\beta, \alpha) n(\alpha, \beta) = 0, 1, 2, \text{ or } 3,
\]

because \( \alpha \) and \( \beta \) are not proportional in \( V \), as in Section 19.9. Suppose for the moment that

\[
(29.1.9) \quad (\alpha, \alpha) \geq (\beta, \beta),
\]

so that

\[
(29.1.10) \quad |n(\beta, \alpha)| \leq |n(\alpha, \beta)|.
\]

This implies that

\[
(29.1.11) \quad |n(\beta, \alpha)| = 0 \text{ or } 1,
\]

by (29.1.8). If \( n(\beta, \alpha) = -1 \), then \( A = 0 \), as desired.

Otherwise, if \( n(\beta, \alpha) = 0 \) or 1, then \( (\beta, \alpha) \geq 0 \). This implies that

\[
(29.1.12) \quad (\beta + \alpha, \beta + \alpha) > (\alpha, \alpha), (\beta, \beta).
\]

Note that \( \alpha + \beta \in \Phi_1 \), because \( \alpha + \beta \in \Phi \), by hypothesis. It follows that

\[
(29.1.13) \quad (\alpha, \alpha) = (\beta, \beta)
\]

in this case, because there are at most two possible values for the norms of elements of \( \Phi_1 \), as before.

Similarly,

\[
(29.1.14) \quad (\beta + 2 \alpha, \beta + 2 \alpha) > (\beta + \alpha, \beta + \alpha),
\]

because \( (\beta, \alpha) \geq 0 \). This implies that \( \beta + 2 \alpha \notin \Phi_1 \), because there are at most two possible values of the norms of elements of \( \Phi_1 \). This means that \( \beta + 2 \alpha \notin \Phi \). Remember that \( \beta + j \alpha \in \Phi \) when \( j \in \mathbb{Z} \) and \( -r < j < q \), as in Section 20.5. It
follows that \( q \leq 1 \), by the definition \( q \). Thus \( q = 1 \), because \( \alpha + \beta \in \Phi \). This implies that \( B = 0 \), because of (29.1.13).

Suppose now that
\[
(\alpha, \alpha) < (\beta, \beta).
\]
Observe that
\[
(\alpha + \beta, \alpha + \beta) = (\alpha, \alpha) \text{ or } (\beta, \beta),
\]
because \( \alpha + \beta \in \Phi_1 \), and there are at most two possible values for the norms of elements of \( \Phi_1 \). This implies that \( (\alpha, \beta) < 0 \), so that
\[
n(\alpha, \beta) < 0.
\]
We also get that
\[
(\beta - \alpha, \beta - \alpha) > (\beta, \beta).
\]
It follows that \( \beta - \alpha \notin \Phi \), because otherwise \( \beta - \alpha \) would be an element of \( \Phi_1 \), and have the same norm as \( \alpha \) or \( \beta \).

This shows that \( r = 0 \), because \( \beta + j \alpha \in \Phi \) when \( j \in \mathbb{Z} \) and \( -r < j < q \), as before. Using (29.1.15) and the fact that \( (\alpha, \beta) \neq 0 \), we obtain that
\[
|n(\alpha, \beta)| < |n(\beta, \alpha)|.
\]
Combining this with (29.1.8), we get that \( n(\alpha, \beta) = -1, 0, \) or \( 1 \). This means that
\[
n(\alpha, \beta) = -1,
\]
because of (29.1.17).

Because \( r = 0 \), (29.1.3) reduces to saying that \( q = -n(\beta, \alpha) \). Equivalently,
\[
q = \frac{n(\beta, \alpha)}{n(\alpha, \beta)},
\]
by (29.1.20). This implies that
\[
q = \frac{(\beta, \beta)}{(\alpha, \alpha)},
\]
by the definition of \( n(\cdot, \cdot) \). It follows that \( B = 0 \) in this case as well.

### 29.2 An identity for root strings

Let \( V \) be a vector space over the real numbers of positive finite dimension, and let \( \Phi \) be a root system in \( V \). Also let \( (\cdot, \cdot) \) be an inner product on \( V \) that is invariant under the Weyl group of \( \Phi \), and suppose that \( \alpha, \beta \in \Phi \) are not proportional in \( V \). As before, we take \( r, q \) to be the largest nonnegative integers such that \( \beta - r \alpha, \beta + q \alpha \in \Phi \). Similarly, let \( r', q' \) be the largest nonnegative integers such that
\[
\alpha - r' \beta, \alpha + q' \beta \in \Phi.
\]
Thus
\[(29.2.2)\]
\[2(\alpha,\beta)(\beta,\beta)^{-1} = r' - q',\]
as in (29.1.3) and Section 20.5. We would like to check that
\[(29.2.3)\]
\[q(r + 1)(\beta,\beta)^{-1} = q'(r' + 1)(\alpha,\alpha)^{-1},\]
as in Exercise 10 on p47 of [13]. This is mentioned in Exercise 3 on p150 of [13] as well.

Remember that \(\beta + j\alpha \in \Phi\) for every integer \(j\) with \(-r \leq j \leq q\), as in Section 20.5. Similarly, if \(l \in \mathbb{Z}\) and \(-r' \leq l \leq q'\), then \(\alpha + l\beta \in \Phi\).

It follows that \(q = 0\) if and only if \(\alpha + \beta \not\in \Phi\). Of course, \(q' = 0\) if and only if \(\alpha + \beta \not\in \Phi\) too, for the same reason. Thus \(q = 0\) if and only if \(q' = 0\), in which case (29.2.3) holds trivially.

Similarly, \(r = 0\) if and only if \(\beta - \alpha \not\in \Phi\). The same argument shows that \(r' = 0\) if and only if \(\alpha - \beta \not\in \Phi\). Of course, \(\alpha - \beta \in \Phi\) if and only if \(\beta - \alpha \in \Phi\). This means that \(r = 0\) if and only if \(r' = 0\). In this case, (29.2.3) follows from (29.1.3) and (29.2.2).

Thus we may suppose from now on that \(q, q', r, r' \geq 1\). Remember that
\[(29.2.4)\]
\[g + r, q' + r' \leq 3,\]
as in Section 20.5.

Observe that \(q = r\) if and only if \((\alpha, \beta) = 0\), by (29.1.3). Similarly, \(q' = r'\) if and only if \((\alpha, \beta) = 0\). Under these conditions, we get that \(q = r = q' = r' = 1\), because of (29.2.4). Thus it suffices to verify that \((\alpha, \alpha) = (\beta, \beta)\), to get (29.2.3).

Remember that \(\alpha + \beta \in \Phi\), because \(q, q' \geq 1\). We also have that
\[(29.2.5)\]
\[(\alpha, \alpha + \beta) = (\alpha, \alpha), \quad (\beta, \alpha + \beta) = (\beta, \beta)\]
and
\[(29.2.6)\]
\[(\alpha + \beta, \alpha + \beta) = (\alpha, \alpha) + (\beta, \beta),\]
because \((\alpha, \beta) = 0\). In particular,
\[(29.2.7)\]
\[
\max((\alpha, \alpha), (\beta, \beta)) < (\alpha + \beta, \alpha + \beta) \leq 2 \max((\alpha, \alpha), (\beta, \beta)).
\]
Note that \(\alpha + \beta\) is not proportional to either \(\alpha\) or \(\beta\) in \(V\).

Using the remarks in Section 19.9, we obtain that \((\alpha + \beta, \alpha + \beta)\) is either 1, 2, or 3 times \((\alpha, \alpha)\), and similarly for \((\beta, \beta)\). In fact, \((\alpha + \beta, \alpha + \beta)\) has to be 2 times \((\alpha, \alpha)\) or \((\beta, \beta)\), because of (29.2.7). In either case, we get that \((\alpha, \alpha) = (\beta, \beta)\), as desired.

Suppose now that \((\alpha, \beta) \neq 0\), so that \(q \neq r\) and \(q' \neq r'\). It follows that
\[(29.2.8)\]
\[|q - r| = |q' - r'| = 1,\]
because of (29.2.4). This implies that \((\alpha, \alpha) = (\beta, \beta)\), because of (29.1.3) and (29.2.2). Using this and (29.1.3), (29.2.2) again, we get that
\[(29.2.9)\]
\[r - q = r' - q'.\]
One can use this and (29.2.4) to verify that \(q = q'\) and \(r = r'\). More precisely, either \(q = q'\) is equal to 1 and \(r = r'\) is equal to 2, or the other way around. This implies (29.2.3), because \((\alpha, \alpha) = (\beta, \beta)\).
29.3 Related properties of $\lambda_\alpha$’s

Let $V$ be a vector space over the real numbers of positive finite dimension again, and let $\Phi$ be a root system in $V$. If $\alpha \in \Phi$, then the symmetry $\sigma_\alpha$ on $V$ with vector $\alpha$ that maps $\Phi$ onto itself can be expressed as $\sigma_\alpha(v) = v - \lambda_\alpha(v) \alpha$, where $\lambda_\alpha$ is a linear functional on $V$ such that $\lambda_\alpha(\alpha) = 2$, as usual. Let $\langle \cdot, \cdot \rangle$ be an inner product on $V$ that is invariant under the Weyl group of $\Phi$, so that $\sigma_\alpha$ is the reflection on $V$ with respect to $\langle \cdot, \cdot \rangle$ associated to $\alpha$ for every $\alpha \in \Phi$. This means that $\lambda_\alpha(v) = 2 \langle v, \alpha \rangle \langle \alpha, \alpha \rangle^{-1}$ for every $\alpha \in \Phi$ and $v \in V$, as before.

Suppose that $\alpha, \beta \in \Phi$ are not proportional in $V$, and let $r, q, r', q' \geq 0$ be as in the previous section. If we multiply both sides of (29.2.3) by $2 \langle \alpha, \beta \rangle$, then we get that

$$q (r + 1) \lambda_\beta(\alpha) = q' (r' + 1) \lambda_\alpha(\beta).$$

(29.3.1)

If $\langle \alpha, \beta \rangle = 0$, then $\lambda_\alpha(\beta) = \lambda_\beta(\alpha) = 0$, $q = q'$, and $r = r'$, as in the previous section.

Suppose now that $\Phi$ is a reduced root system, and that $\alpha + \beta \in \Phi$ too. If $v \in V$, then

$$\lambda_{\alpha+\beta}(v) = 2 \langle \alpha + \beta, \alpha + \beta \rangle^{-1} (v, \alpha + \beta)$$

$$= (\alpha + \beta, \alpha + \beta)^{-1} (\alpha, \alpha) \lambda_\alpha(v) + (\alpha + \beta, \alpha + \beta)^{-1} (\beta, \beta) \lambda_\beta(v).$$

(29.3.2)

It follows that

$$\lambda_{\alpha+\beta}(v) = q' (r' + 1)^{-1} \lambda_\alpha(v) + q (r + 1)^{-1} \lambda_\beta(v)$$

for every $v \in V$, using (29.1.4) and its analogue for $q', r'$. Note that $\lambda_\alpha$ and $\lambda_\beta$ are not proportional as linear functionals on $V$, because $\alpha$ and $\beta$ are not proportional in $V$. Thus the coefficients of $\lambda_\alpha$ and $\lambda_\beta$ on the right side of (29.3.3) are uniquely determined by this condition.

29.4 Associated Lie algebras

Let $V$ be a vector space over the real numbers of positive finite dimension again, and let $\Phi$ be a reduced root system in $V$. If $\alpha \in \Phi$, then the symmetry $\sigma_\alpha$ on $V$ with vector $\alpha$ that maps $\Phi$ onto itself can be expressed as $\sigma_\alpha(v) = v - \lambda_\alpha(v) \alpha$, where $\lambda_\alpha$ is a linear functional on $V$ with $\lambda_\alpha(\alpha) = 2$. If $\alpha, \beta \in \Phi$, then

$$n(\beta, \alpha) = \lambda_\alpha(\beta) \in \mathbb{Z},$$

(29.4.1)

by the definition of a root system. Of course, this is the same as (29.1.1), where $\langle \cdot, \cdot \rangle$ is an inner product on $V$ that is invariant under the Weyl group. Let $\Theta$ be the subgroup of $V$, as a commutative group with respect to addition, that is generated by $\Phi$.

Let $k$ be a field of characteristic 0. We would like to consider a Lie algebra $(A, [\cdot, \cdot]_A)$ over $k$ as in Serre’s theorem in Section 27.8, or equivalently as in Section 28.7. Thus, if $\alpha \in \Phi \cup \{0\}$, then we ask that $A_\alpha$ be a linear subspace
of $A$, and that $A$ correspond to the direct sum of $A_\alpha$, $\alpha \in \Phi \cup \{0\}$, as a vector space over $k$. Put $A_\alpha = \{0\}$ when $\alpha \in \Theta \setminus (\Phi \cup \{0\})$, so that we may ask that $[A_\alpha, A_\beta] \subseteq A_{\alpha+\beta}$ for every $\alpha, \beta \in \Theta$. We also ask that $A_\alpha$ have dimension one as a vector space over $k$ when $\alpha \in \Phi$.

As usual, $A_\alpha'$ denotes the dual space of linear functionals on $A_\alpha$, as a vector space over $k$. We ask that there be a group homomorphism $\phi : \Theta \to A_\alpha'$, as a commutative group with respect to addition, such that $\text{ad}_w(x) = \phi_\alpha(w)x$ for every $w \in A_\alpha$, $x \in A_\alpha$, and $\alpha \in \Phi \cup \{0\}$, where $\phi_\alpha = \phi(\alpha)$. Of course, this implies that $A_0$ is a Lie subalgebra of $A$ that is commutative as a Lie algebra.

If $\alpha, \beta \in \Phi$ and $\alpha + \beta \in \Phi$, then we ask that

\begin{equation}
[A_\alpha, A_\beta] = A_{\alpha+\beta}.
\end{equation}

We also ask that for each $\alpha \in \Phi$ there be an element $h_\alpha$ of $[A_\alpha, A_{-\alpha}] \subseteq A_0$ such that

\begin{equation}
\phi_\beta(h_\alpha) = \lambda_\alpha(\beta)
\end{equation}

for every $\beta \in \Phi$. In particular, this means that $\phi_\alpha(h_\alpha) = 2 \cdot 1 = 1 + 1$ in $k$, which is not zero, because $k$ has characteristic zero. Of course, $[A_\alpha, A_{-\alpha}]$ has dimension at most one as a vector space over $k$, because $A_\alpha$, $A_{-\alpha}$ are one-dimensional. It follows that $[A_\alpha, A_{-\alpha}]$ is the one-dimensional linear subspace of $A$ spanned by $h_\alpha$, and that $h_\alpha$ is uniquely determined by (29.4.3), with $\beta = \alpha$.

Let $\Delta$ be a base for $\Phi$. We ask that $h_\alpha$, $\alpha \in \Delta$, be a basis for $A_0$, as a vector space over $k$. If $A$ is as in Section 22.1, then conditions like these were discussed in Section 22.11. We also have that $A$ is isomorphic to the Lie algebra obtained as in Serre’s theorem in Section 27.8, as discussed in Section 28.7. If $A$ is obtained as in Serre’s theorem, then conditions like these follow from properties discussed in Sections 27.11, 27.15, 28.2, and 28.10.

If $\beta \in \Phi$, then (29.4.3) defines a $k$-valued function of $\alpha \in \Delta$, which is the same as (27.3.9) when $\beta \in \Delta$. The collection of these functions with $\beta \in \Delta$ is a basis for the space of $k$-valued functions on $\Delta$, as a vector space over $k$, as in Section 27.3. This means that $\phi_\beta$, $\beta \in \Delta$, is a basis for $A_\alpha'$ as a vector space over $k$. If $A$ is as in Section 22.1, then this can be obtained from the facts that $A_\alpha'$ is spanned by $\phi_\beta$, $\beta \in \Delta$, as a vector space over $k$, and the dimension of $A_\alpha'$ is the same as the number of elements of $\Delta$, which is the dimension of $V$. If $h \in A_0$, then it follows that $h$ is uniquely determined by $\phi_\beta(h)$, $\beta \in \Delta$.

Remember that $\Theta$ is the same as the subgroup of $V$, as a commutative group with respect to addition, generated by $\Delta$. The remark in the preceding paragraph implies that $\phi$ is injective as a mapping from $\Theta$ into $A_\alpha'$. This also uses the hypothesis that $k$ have characteristic $0$. If $\alpha \in \Phi \cup \{0\}$, then one can use the previous statement to get that $A_\alpha$ consists of all $x \in A$ such that $\text{ad}_w(x) = \phi_\alpha(w)x$ for every $w \in A_0$.

If $\alpha \in \Phi$, then $\lambda_\alpha$ is an element of the dual $V'$ of $V$, as a vector space over $\mathbb{R}$. The set $\Phi'$ of $\lambda_\alpha$, $\alpha \in \Phi$, is a root system in $V'$, as in Section 19.8. The set $\Delta'$ of $\lambda_\alpha$, $\alpha \in \Delta$, is a base for $\Phi'$, as in Section 19.13, and because $\Phi$ is reduced, by hypothesis. Thus, if $\gamma \in \Phi$, then $\lambda_\gamma$ can be expressed as a linear combination of $\lambda_\alpha$, $\alpha \in \Delta$, with coefficients in $\mathbb{Z}$. 


29.5. A LEMMA ABOUT BRACKETS

If \( \gamma \in \Phi \), then it follows that \( h_\gamma \) can be expressed as a linear combination of \( h_\alpha, \alpha \in \Delta \), with the same coefficients in \( \mathbb{Z} \) as in the preceding paragraph. This uses (29.4.3), and the fact that \( h_\gamma \) is uniquely determined by \( \phi_\beta(h_\gamma), \beta \in \Delta \), as before.

29.5 A lemma about brackets

Let us continue with the same notation and hypotheses as in the previous section. Also let \( \alpha, \beta \in \Phi \) be given, and suppose that \( \alpha, \beta \) are not proportional in \( V \). Let \( r, q \) be the largest nonnegative integers such that \( \beta - r \alpha \) and \( \beta + q \alpha \) are elements of \( \Phi \), as before. It follows that

\[
\beta + j \alpha \in \Phi
\]

for every \( j \in \mathbb{Z} \) with \(-r \leq j \leq q\), as in Section 20.5.

Suppose that \( x_\alpha \in A_\alpha \) and \( x_{-\alpha} \in A_{-\alpha} \) satisfy

\[
[x_\alpha, x_{-\alpha}]_A = h_\alpha.
\]

Of course, this implies that \( x_\alpha, x_{-\alpha} \neq 0 \), and (29.5.2) can always be arranged by multiplying \( x_\alpha \) or \( x_{-\alpha} \) by a suitable element of \( k \). If \( x_\beta \in A_\beta \), then

\[
[x_{-\alpha}, [x_\alpha, x_\beta]]_A = q (r + 1) x_\beta.
\]

This corresponds to the lemma on p146 of [13].

Note that \( [x_\alpha, x_\beta]_A \in A_{\alpha + \beta} \), so that the left side of (29.5.3) is an element of \( A_\beta \). If \( \alpha + \beta \notin \Phi \), then \( A_{\alpha + \beta} = \{0\} \), which means that \( [x_\alpha, x_\beta]_A = 0 \). We also have that \( q = 0 \) in this case, so that (29.5.3) holds.

Let \( E \) be the linear subspace of \( A \) spanned by \( A_{\beta + j \alpha}, j \in \mathbb{Z} \). More precisely, we may as well restrict our attention to \(-r \leq j \leq q\), because otherwise \( \beta + j \alpha \) is not in \( \Phi \), and \( A_{\beta + j \alpha} = \{0\} \). We also have that \( E \) corresponds to the direct sum of \( A_{\beta + j \alpha}, -r \leq j \leq q \), because \( A \) corresponds to the direct sum of \( A_\gamma, \gamma \in \Phi \cup \{0\} \), as a vector space over \( k \). It follows that

\[
\dim E = q + r + 1
\]

as a vector space over \( k \), because \( A_\gamma \) is one-dimensional for every \( \gamma \in \Phi \).

The linear span of \( x_\alpha, x_{-\alpha} \), and \( h_\alpha \) in \( A \) is a Lie subalgebra of \( A \) that is isomorphic to \( \mathfrak{sl}_2(k) \). If we consider \( A \) as a module over this subalgebra, then \( E \) is a submodule of \( A \). This is basically the same as in Section 17.9, with \( y_\alpha = x_{-\alpha} \).

The weights of \( h_\alpha \) on \( E \) are given by

\[
\phi_{\beta + j \alpha}(h_\alpha) = \phi_\beta(h_\alpha) + j \phi_\alpha(h_\alpha) = \phi_\beta(h_\alpha) + 2 j,
\]

with \(-r \leq j \leq q\). One can verify that \( E \) is irreducible as a module over \( \mathfrak{sl}_2(k) \), as in Section 17.9.

Under these conditions, \( E \) is as in Section 15.3 as a module over \( \mathfrak{sl}_2(k) \), with \( m = q + r \), as in Section 17.9. To get (29.5.3), one can consider \( v_q \) in Section 15.3, as on p146 of [13].
29.6 Some remarks about automorphisms

Let us continue with the same notation and hypotheses as in Section 29.4. Suppose that \( \sigma \) is a Lie algebra automorphism of \( A \) such that

\[
\sigma(w) = -w
\]

for every \( w \in A_0 \). Let \( w \in A_0 \) and \( x \in A \) be given, and observe that

\[
[w, \sigma(x)]_A = -[\sigma(w), \sigma(x)]_A = -\sigma([w, x]_A).
\]

Also let \( \alpha \in \Phi \) be given, and note that \( \phi_{-\alpha}(w) = -\phi_\alpha(w) \). It follows that \([w, x]_A = \phi_\alpha(w) x \) if and only if

\[
[w, \sigma(x)]_A = \phi_{-\alpha}(w) \sigma(x).
\]

This implies that

\[
\sigma(A_\alpha) = A_{-\alpha},
\]

because \( A_\alpha \) consists of all \( x \in A \) such that \( \text{ad}_w(x) = \phi_\alpha(w) x \) for every \( w \in A_0 \), as in Section 29.4, and similarly for \( -\alpha \). Of course, (29.6.4) holds for \( \alpha = 0 \) too, by hypothesis.

Let \( x_\alpha \in A_\alpha \) and \( y_\alpha \in A_{-\alpha} \) be given, with \( x_\alpha, y_\alpha \neq 0 \). Observe that there are \( \xi_\alpha, \eta_\alpha \in k \) such that

\[
\sigma(x_\alpha) = \xi_\alpha y_\alpha, \quad \sigma(y_\alpha) = \eta_\alpha x_\alpha,
\]

because of (29.6.4) and the analogous statement for \( -\alpha \). This also uses the fact that \( A_\alpha, A_{-\alpha} \) are one-dimensional linear subspaces of \( A \), and we have that \( \xi_\alpha, \eta_\alpha \neq 0 \), because \( \sigma \) is injective on \( A \). It follows that

\[
\sigma([x_\alpha, y_\alpha]_A) = [\sigma(x_\alpha), \sigma(y_\alpha)]_A = -\xi_\alpha \eta_\alpha [x_\alpha, y_\alpha]_A.
\]

Remember that \([x_\alpha, y_\alpha]_A \in A_0 \), so that

\[
\sigma([x_\alpha, y_\alpha]_A) = -[x_\alpha, y_\alpha]_A,
\]

by (29.6.1).

Note that \([x_\alpha, y_\alpha]_A \neq 0 \), because \([A_\alpha, A_{-\alpha}]_A \) is one-dimensional as a linear subspace of \( A \). Thus we get that

\[
\xi_\alpha \eta_\alpha = 1
\]

in \( k \). This implies that

\[
\sigma(x_\alpha) = x_\alpha, \quad \sigma(y_\alpha) = y_\alpha.
\]

It follows that \( \sigma \circ \sigma \) is the identity mapping on \( A \), because \( \alpha \in \Phi \) is arbitrary, and using (29.6.1) on \( A_0 \). Remember that an automorphism on \( A \) with these properties was discussed in Section 28.6.
Let $\alpha \in \Phi$ be given again, and let $x_\alpha$ be a nonzero element of $A_\alpha$. Also let $\zeta_\alpha$ be a nonzero element of $k$, and put

\begin{equation}
(29.6.10)\quad z_\alpha = \zeta_\alpha x_\alpha, \quad z_{-\alpha} = -\sigma(z_\alpha) = -\zeta_\alpha \sigma(x_\alpha).
\end{equation}

Thus

\begin{equation}
(29.6.11)\quad [z_\alpha, z_{-\alpha}]_A = -\zeta_\alpha^2 [x_\alpha, \sigma(x_\alpha)]_A,
\end{equation}

and $[x_\alpha, \sigma(x_\alpha)]_A \neq 0$, as before. If $k$ is algebraically closed, then we can choose $\zeta_\alpha$ so that

\begin{equation}
(29.6.12)\quad [z_\alpha, z_{-\alpha}]_A = h_\alpha.
\end{equation}

Note that $-\zeta_\alpha$ would also work, and that only these two elements of $k$ would have this property.

29.7 Diagonal automorphisms

We continue with the same notation and hypotheses as in Section 29.4. Suppose now that $\tau$ is a Lie algebra automorphism of $A$ such that

\begin{equation}
(29.7.1)\quad \tau(w) = w
\end{equation}

for every $w \in A_0$. If $w \in A_0$ and $x \in A$, then

\begin{equation}
(29.7.2)\quad [w, \tau(x)]_A = [\tau(w), \tau(x)]_A = \tau([w, x]_A).
\end{equation}

If $\alpha \in \Phi$, then it follows that $[w, x]_A = \phi_\alpha(w) x$ if and only if

\begin{equation}
(29.7.3)\quad [w, \tau(x)]_A = \phi_\alpha(w) \tau(x).
\end{equation}

This means that

\begin{equation}
(29.7.4)\quad \tau(A_\alpha) = A_\alpha
\end{equation}

for every $\alpha \in \Phi$, as before.

In this case, $\tau$ is said to be a diagonal automorphism of $A$, as on p87 of [13]. If $\alpha \in \Phi$, then there are nonzero elements $\tau_\alpha$, $\tau_{-\alpha}$ such that $\tau$ corresponds to multiplication by $\tau_\alpha$, $\tau_{-\alpha}$ on $A_\alpha$, $A_{-\alpha}$, respectively, because $A_\alpha$, $A_{-\alpha}$ are one-dimensional linear subspaces of $A$. It is easy to see that

\begin{equation}
(29.7.5)\quad \tau_\alpha \tau_{-\alpha} = 1,
\end{equation}

because $[A_\alpha, A_{-\alpha}]$ is a nontrivial linear subspace of $A_0$.

Of course, any Lie algebra automorphism of $A$ is uniquely determined by its values on a set of generators of $A$ as a Lie algebra over $k$. If $A$ is as in Serre’s theorem in Section 27.8, then the diagonal automorphisms of $A$ can be obtained as in Section 28.6.

Let $\sigma$ be a Lie algebra automorphism of $A$ that satisfies (29.6.1), and thus (29.6.4) for every $\alpha \in \Phi$. If $\tau$ is a diagonal automorphism of $A$, then one can check that

\begin{equation}
(29.7.6)\quad \sigma \circ \tau = \tau^{-1} \circ \sigma.
\end{equation}
Observe that $\sigma \circ \tau$ is a Lie algebra automorphism of $A$ that sends every $w \in A_0$ to $-w$. If $\tilde{\sigma}$ is any Lie algebra automorphism of $A$ that sends every $w \in A_0$ to $-w$, then
\begin{equation}
\sigma^{-1} \circ \tilde{\sigma}
\end{equation}
is a diagonal automorphism on $A$.

Let $\tau$ be a diagonal automorphism on $A$ again, and for each $\alpha \in \Phi$, let $\tau_\alpha$ be the nonzero element of $k$ such that $\tau$ corresponds to multiplication by $\tau_\alpha$ on $A_\alpha$, as before. Similarly, we may as well put $\tau_0 = 1$. If $\alpha, \beta \in \Phi$ and $\alpha + \beta \in \Phi$, then it is easy to see that
\begin{equation}
\tau_{\alpha + \beta} = \tau_\alpha \tau_\beta,
\end{equation}
because $[A_\alpha, A_\beta] = A_{\alpha + \beta} \neq \{0\}$.

Let $\Delta$ be a base for $\Phi$, and let $\beta \in \Phi$ be given. We would like to check that $\tau_\beta$ is uniquely determined by the values of $\tau_\alpha$ with $\alpha \in \Delta$. Suppose first that $\beta$ is a positive root with respect to $\Delta$, so that $\beta$ can be expressed as a linear combination of elements of $\Delta$ with nonnegative coefficients. Remember that $\beta$ can be expressed $\sum_{j=1}^{r} \alpha_j$, where $\alpha_j \in \Delta$ for every $j = 1, \ldots, r$, and $\sum_{j=1}^{r} \alpha_j \in \Phi$ for every $l = 1, \ldots, r$, as in Section 19.12. This implies that
\begin{equation}
\tau_\beta = \tau_{\alpha_1} \cdots \tau_{\alpha_r},
\end{equation}
by (29.7.8). Otherwise, if $-\beta$ is a positive root, then one can reduce to the previous case, using (29.7.5).

Let $\Theta$ be the subgroup of $V$, as a commutative group with respect to addition, generated by $\Phi$, which is the same as the subgroup of $V$ generated by $\Delta$. Every mapping from $\Delta$ into $k \setminus \{0\}$ has a unique extension to a group homomorphism from $\Theta$ into the multiplicative group of nonzero elements of $k$. In particular, $\alpha \mapsto \tau_\alpha$, $\alpha \in \Delta$, has a unique extension to a group homomorphism from $\Theta$ into $k \setminus \{0\}$. If $\beta \in \Phi$, then $\tau_\beta$ is the same as the value of this extension at $\beta$, by the remarks in the preceding paragraph. Conversely, every group homomorphism from $\Theta$ into $k \setminus \{0\}$ corresponds to a diagonal automorphism on $A$ in this way.

### 29.8 Basis vectors in $A$

Let us continue with the same notation and hypotheses as in Section 29.4. As before, we let $\Delta$ be a base for $\Phi$, and ask that $h_\alpha$, $\alpha \in \Delta$, be a basis for $A_0$, as a vector space over $k$. Let $z_\alpha$ be a nonzero element of $A_\alpha$ for each $\alpha \in \Phi$. Note that the collection of $z_\alpha$'s, $\alpha \in \Phi$, together with $h_\alpha$, $\alpha \in \Delta$, is a basis for $A$ as a vector space over $k$.

If $\alpha, \beta \in \Phi$ and $\alpha + \beta \in \Phi$, then there is a unique element $c_{\alpha, \beta}$ of $k$ such that
\begin{equation}
[z_\alpha, z_\beta]_A = c_{\alpha, \beta} z_{\alpha + \beta},
\end{equation}
because $[z_\alpha, z_\beta]_A$ is an element of $A_{\alpha + \beta}$, which is spanned by $z_{\alpha + \beta}$. Of course, if $\alpha, \beta \in \Phi$ and $\alpha + \beta \notin \Phi \cup \{0\}$, then $[z_\alpha, z_\beta]_A = 0$, because $[z_\alpha, z_\beta] \in A_{\alpha + \beta} = \{0\}$. 

Observe that
\[(29.8.2) \quad [h_\alpha, z_\beta]_A = \phi_\beta(h_\alpha) z_\beta = \lambda_\alpha(\beta) z_\beta\]
for every \(\alpha, \beta \in \Phi\), as in Section 29.4. If \(\gamma \in \Phi\), then \(h_\gamma\) can be expressed as a linear combination of \(h_\alpha, \alpha \in \Delta\), with coefficients in \(\mathbb{Z}\) obtained from the dual root system \(\Phi'\), as in Section 29.4.

As in Section 28.6, there is a Lie algebra automorphism \(\sigma\) on \(A\) such that \(\sigma(w) = -w\) for every \(w \in A_0\), and \(\sigma \circ \sigma\) is the identity mapping on \(A\). It follows that \(\sigma(A_\alpha) = A_{-\alpha}\) for every \(\alpha \in \Phi\), as in Section 29.6. Let \(\Phi^+\) be the set of positive roots with respect to \(\Delta\), which is to say the set of \(\alpha \in \Phi\) that can be expressed as linear combinations of elements of \(\Delta\) with nonnegative coefficients.

Suppose for the moment that \(k\) is algebraically closed. If \(\alpha \in \Phi^+\), then there is a nonzero element \(z_\alpha\) of \(A_\alpha\) such that
\[(29.8.3) \quad z_{-\alpha} = -\sigma(z_\alpha)\]
satisfies
\[(29.8.4) \quad [z_\alpha, z_{-\alpha}]_A = h_\alpha,\]
as in Section 29.6. Note that \(z_{-\alpha}\) is a nonzero element of \(A_{-\alpha}\). By doing this for every \(\alpha \in \Phi^+\), we get a nonzero element \(z_\alpha\) of \(A_\alpha\) for every \(\alpha \in \Phi\). We also have that the analogues of (29.8.3) and (29.8.4) hold with \(\alpha\) replaced with \(-\alpha\), because \(\sigma \circ \sigma\) is the identity mapping on \(A\), and \(h_{-\alpha} = -h_\alpha\).

Let \(z_\alpha\) be a nonzero element of \(A_\alpha\) for each \(\alpha \in \Phi\) again, and suppose that (29.8.4) holds for every \(\alpha \in \Phi\). Also let \(c_{\alpha, \beta} \in k\) be defined for \(\alpha, \beta \in \Phi\) with \(\alpha + \beta \in \Phi\) as in (29.8.1), as before. Note that \([\cdot, \cdot]_A\) is uniquely determined on \(A\) by the \(c_{\alpha, \beta}\)'s under these conditions.

Suppose that (29.8.3) holds for every \(\alpha \in \Phi\). If \(\alpha, \beta \in \Phi\) and \(\alpha + \beta \in \Phi\), then
\[(29.8.5) \quad [z_{-\alpha}, z_{-\beta}]_A = [-\sigma(z_\alpha), -\sigma(z_\beta)]_A = \sigma([z_\alpha, z_\beta]_A) = c_{\alpha, \beta} \sigma(z_{\alpha + \beta}) = -c_{\alpha, \beta} z_{-\alpha - \beta}.\]
Of course, \(-\alpha, -\beta, -\alpha - \beta \in \Phi\) in this case, so that
\[(29.8.6) \quad [z_{-\alpha}, z_{-\beta}]_A = c_{-\alpha, -\beta} z_{-\alpha - \beta},\]
as in (29.8.1). It follows that
\[(29.8.7) \quad c_{-\alpha, -\beta} = -c_{\alpha, \beta}\]
for every \(\alpha, \beta \in \Phi\) with \(\alpha + \beta \in \Phi\). This corresponds to Theorem 10 on p51 of [23], and to parts (a) and (b) of the proposition on p146 of [13].

Let \(z_\alpha\) be a nonzero element of \(A_\alpha\) for each \(\alpha \in \Phi\) again, and suppose that (29.8.4) holds for every \(\alpha \in \Phi\). If \(\alpha, \beta \in \Phi\) and \(\alpha + \beta \in \Phi\), then let \(c_{\alpha, \beta} \in k\) be as in (29.8.1), and suppose that (29.8.7) holds. Under these conditions, the collection of \(z_\alpha\)'s, \(\alpha \in \Phi\), together with the \(h_\alpha\)'s, \(\alpha \in \Delta\), is said to be a Chevalley basis for \(A\), as on p147 of [13].
29.9 Some properties of the $c_{\alpha, \beta}$’s

We continue with the same notation and hypotheses as in Section 29.4. Let $z_\alpha$ be a nonzero element of $A_\alpha$ for every $\alpha \in \Phi$ again, and suppose that (29.8.4) holds for every $\alpha \in \Phi$. If $\alpha, \beta \in \Phi$ and $\alpha + \beta \in \Phi$, then let $c_{\alpha, \beta} \in k$ be as in (29.8.1).

Let $\alpha, \beta \in \Phi$ with $\alpha + \beta \in \Phi$ be given, so that $-\alpha, -\beta, -\alpha - \beta \in \Phi$ too. Observe that

\[(29.9.1) \quad [z_\alpha, z_\beta]_A, [z_\alpha, z_\beta]_A \quad = \quad [c_{\alpha, \beta} z_{\alpha + \beta}, c_{-\alpha, -\beta} z_{-\alpha - \beta}]_A \quad = \quad c_{\alpha, \beta} c_{-\alpha, -\beta} h_{\alpha + \beta}.\]

We can also use the Jacobi identity to get that

\[(29.9.2) \quad [z_\alpha, z_\beta]_A, [z_\alpha, z_\beta]_A \quad = \quad [z_\beta, [z_\alpha, z_\beta]_A]_A - [z_\alpha, [z_\beta, z_\alpha]_A]_A - [z_\beta, [z_\alpha, z_\beta]_A]_A.\]

Note that $\alpha$ and $\beta$ are not proportional in $V$, because $\alpha + \beta \in \Phi$, and $\Phi$ is reduced. Let $r, q$ be the largest nonnegative integers such that $\beta - r \alpha$ and $\beta + q \alpha$ are elements of $\Phi$ again. Similarly, let $r', q'$ be the largest nonnegative integers such that $\alpha - r' \beta$ and $\alpha + q' \beta$ are elements of $\Phi$. Equivalently, $r, q$ are the largest nonnegative integers such that

\[(29.9.3) \quad (-\beta) - r (-\alpha), (-\beta) + q (-\alpha) \in \Phi,\]

and $r', q'$ are the largest nonnegative integers such that

\[(29.9.4) \quad (-\alpha) - r' (-\beta), (-\alpha) + q' (-\beta) \in \Phi.\]

Of course, $q, q' \geq 1$, because $\alpha + \beta \in \Phi$.

Note that

\[(29.9.6) \quad [z_\alpha, [z_\alpha, z_\beta]_A]_A = q (r + 1) z_{-\beta},\]

as in (29.5.3). Similarly,

\[(29.9.7) \quad [z_\beta, [z_\beta, z_\alpha]_A]_A = q' (r' + 1) z_{-\alpha}.\]

Combining these identities with (29.9.2), we get that

\[(29.9.8) \quad [[z_\alpha, z_\beta]_A, [z_\alpha, z_\beta]_A]_A = -q' (r' + 1) [z_\alpha, z_\alpha]_A - q (r + 1) [z_\beta, z_\beta]_A = -q' (r' + 1) h_\alpha - q (r + 1) h_\beta.\]

It follows that

\[(29.9.9) \quad c_{\alpha, \beta} c_{-\alpha, -\beta} h_{\alpha + \beta} = -q' (r' + 1) h_\alpha - q (r + 1) h_\beta,\]
29.10. **UNIQUENESS OF CHEVALLEY BASES**

by (29.9.1).

If $\gamma \in \Phi$, then we can evaluate $\phi_\gamma$ at both sides of (29.9.9), using (29.4.3), to get

$$c_{\alpha,\beta} c_{-\alpha,-\beta} \lambda_{\alpha+\beta}(\gamma) = -q'(r' + 1) \lambda_\alpha(\gamma) - q(r + 1) \lambda_\beta(\gamma).$$

In particular, we can take $\gamma = \alpha + \beta$, to get that $2 c_{\alpha,\beta} c_{-\alpha,-\beta}$ corresponds to an integer under the natural embedding of $Q$ into $k$. This means that

$$c_{\alpha,\beta} c_{-\alpha,-\beta} \lambda_{\alpha+\beta} = -q'(r' + 1) \lambda_\alpha - q(r + 1) \lambda_\beta$$

as linear functionals on $V$, because $V$ is spanned by $\Phi$. Comparing this with (29.3.3), we obtain that

$$c_{\alpha,\beta} c_{-\alpha,-\beta} = -(r + 1)^2 = -(r' + 1)^2.$$

This also uses (29.9.5), and corresponds to a remark on p148 of [13].

If (29.8.7) holds, then we get that

$$c_{\alpha,\beta}^2 = (r + 1)^2 = (r' + 1)^2.$$

This means that

$$c_{\alpha,\beta} = \pm (r + 1) = \pm (r' + 1),$$

which is part of a famous theorem of Chevalley, as in part (d) of the theorem on p147 of [13], and Theorem 11 on p51 of [23]. More precisely, some of the previous arguments correspond to the proof of part (c) of the proposition on p146 of [13].

Remember $\Delta$ is a base for $\Phi$, and that the $z_\alpha$'s, $\alpha \in \Phi$, together with the $h_\alpha$'s, $\alpha \in \Delta$, form a basis for $A$ as a vector space over $k$, which is a Chevalley basis, as in the previous section. Observe that the brackets of any two elements of this basis can be expressed as a linear combination of elements of the basis with integer coefficients, because of (29.9.14). This corresponds to the theorem on p147 of [13].

Let $A(\mathbb{Z})$ be the collection of elements of $A$ that can be expressed as linear combinations of the $z_\alpha$'s, $\alpha \in \Phi$, and the $h_\alpha$'s, $\alpha \in \Delta$, with integer coefficients. This is a subgroup of $A$, as a commutative group with respect to addition, and in fact a Lie subalgebra of $A$, as a Lie algebra over $\mathbb{Z}$. This is part of Remark (1) on p51 of [23], and is also mentioned on p149 of [13].

Note that $h_\gamma \in A(\mathbb{Z})$ for every $\gamma \in \Phi$, as in Section 29.4. This implies that $A(\mathbb{Z})$ does not depend on the base $\Delta$ for $\Phi$, as mentioned on p149 of [13].

### 29.10 Uniqueness of Chevalley bases

Let us continue with the same notation and hypotheses as in Section 29.4. Let $z_\alpha$, $\tilde{z}_\alpha$ be nonzero elements of $A_\alpha$ for every $\alpha \in \Phi$. Thus, for each $\alpha \in \Phi$, there is a nonzero element $\eta_\alpha$ of $k$ such that

$$\tilde{z}_\alpha = \eta_\alpha z_\alpha.$$
Suppose that \([\bar{z}, z_\alpha]_A = h_\alpha\) for every \(\alpha \in \Phi\), as before. Observe that

\begin{equation}
[\bar{z}, z_\alpha]_A = h_\alpha
\end{equation}

for every \(\alpha \in \Phi\) if and only if

\begin{equation}
\eta_\alpha \eta_{-\alpha} = 1
\end{equation}

for every \(\alpha \in \Phi\), as on p148 of [13].

If \(\alpha, \beta \in \Phi\) and \(\alpha + \beta \in \Phi\), then there is a unique \(c_{\alpha, \beta} \in k\) such that

\begin{equation}
[\bar{z}_\alpha, \bar{z}_\beta]_A = c_{\alpha, \beta} z_{\alpha + \beta},
\end{equation}

as before. Similarly, there is a unique \(\bar{c}_{\alpha, \beta} \in k\) such that

\begin{equation}
[\bar{z}_\alpha, \bar{z}_\beta]_A = \bar{c}_{\alpha, \beta} \bar{z}_{\alpha + \beta}.
\end{equation}

More precisely, \(c_{\alpha, \beta}, \bar{c}_{\alpha, \beta} \neq 0\), because of (29.4.2). Using (29.10.1), we get that

\begin{equation}
\bar{c}_{\alpha, \beta} \eta_{\alpha + \beta} = \eta_\alpha \eta_\beta c_{\alpha, \beta}.
\end{equation}

Suppose that \(c_{\alpha, \beta} = c_{\alpha, \beta}^\prime\) for every \(\alpha, \beta \in \Phi\) with \(\alpha + \beta \in \Phi\), as before. Suppose also that (29.10.3) holds for every \(\alpha \in \Phi\). Under these conditions, one can check that

\begin{equation}
\bar{c}_{\alpha, \beta} = -\bar{c}_{\alpha, \beta}
\end{equation}

for every \(\alpha, \beta \in \Phi\) with \(\alpha + \beta \in \Phi\) if and only if

\begin{equation}
\eta_{\alpha + \beta} = \pm \eta_\alpha \eta_\beta
\end{equation}

for every \(\alpha, \beta \in \Phi\) with \(\alpha + \beta \in \Phi\), as on p148 of [13].

Let \(\Delta\) be a base for \(\Phi\), and remember that \(\Theta\) is the subgroup of \(V\), as a commutative group with respect to addition, generated by \(\Phi\). This is the same as the subgroup of \(V\) generated by \(\Delta\), and every mapping from \(\Delta\) into \(k \setminus \{0\}\) has a unique extension to a group homomorphism from \(\Theta\) into the multiplicative group of nonzero elements of \(k\).

Suppose now that \(\eta_\alpha\) is a nonzero element of \(k\) for each \(\alpha \in \Phi\) that satisfies (29.10.3) for every \(\alpha \in \Phi\), and (29.10.7) for every \(\alpha, \beta \in \Phi\) with \(\alpha + \beta \in \Phi\). Let \(\alpha \mapsto \tau_\alpha\) be the unique group homomorphism from \(\Theta\) into the multiplicative group of nonzero elements of \(k\) such that

\begin{equation}
\tau_\alpha = \eta_\alpha \quad \text{for every } \alpha \in \Delta.
\end{equation}

If \(\alpha \in \Phi\), then let \(\rho_\alpha\) be the nonzero element of \(k\) determined by

\begin{equation}
\eta_\alpha = \rho_\alpha \tau_\alpha.
\end{equation}

Observe that

\begin{equation}
\rho_\alpha \rho_{-\alpha} = 1
\end{equation}

for every \(\alpha \in \Phi\), and that

\begin{equation}
\rho_{\alpha + \beta} = \pm \rho_\alpha \rho_\beta
\end{equation}
for every $\alpha, \beta \in \Phi$ with $\alpha + \beta \in \Phi$, because of the analogous properties of the $\eta_\alpha$’s and $\tau_\alpha$’s. We would like to verify that

$$(29.10.12) \quad \rho_\beta = \pm 1$$

for every $\beta \in \Phi$.

Of course, $\rho_\alpha = 1$ for every $\alpha \in \Delta$, by construction. Suppose that $\beta \in \Phi$ is a positive root with respect to $\Delta$, so that $\beta$ can be expressed as a linear combination of elements of $\Delta$ with nonnegative coefficients. In this case, $\beta$ can be expressed as $\sum_{j=1}^{r} \alpha_j$, where $\alpha_j \in \Delta$ for every $j = 1, \ldots, r$, and $\sum_{j=1}^{l} \alpha_j \in \Phi$ for every $l = 1, \ldots, r$, as in Section 19.12. This permits one to obtain (29.10.12) using (29.10.11). If $-\beta$ is a positive root, then (29.10.12) can be obtained from (29.10.10) and the previous case.

Let the $z_\alpha$’s and $e_\alpha$’s be as before, so that the $z_\alpha$’s, $\alpha \in \Phi$, together with the $h_\alpha$’s, $\alpha \in \Delta$, is a Chevalley basis for $A$, and similarly for the $\tilde{z}_\alpha$’s. The corresponding $\eta_\alpha$’s can be defined for $\alpha \in \Phi$ as in (29.10.1), and satisfy (29.10.3) and (29.10.7). Thus we can define $\tau_\alpha$ and $\rho_\alpha$ as in (29.10.8) and (29.10.9), respectively. Using the $\tau_\alpha$’s, we get a diagonal automorphism $\tau$ of $A$, as in Section 29.7. Note that $\tau(h) = h$ for every $h \in A_0$, and $\tau(z_\alpha) = \tau_\alpha z_\alpha$ for every $\alpha \in \Phi$, as before.

If $\alpha \in \Phi$, then we get that

$$(29.11.1) \quad \tilde{z}_\alpha = \rho_\alpha \tau(z_\alpha).$$

Let $A(\mathbb{Z})$ be the collection of elements of $A$ that can be expressed as linear combinations of the $z_\alpha$’s, $\alpha \in \Phi$, and $h_\alpha$’s, $\alpha \in \Delta$, with coefficients in $\mathbb{Z}$, as in the previous section. Similarly, let $\tilde{A}(\mathbb{Z})$ be the collection of elements of $\tilde{A}$ that can be expressed as linear combinations of the $\tilde{z}_\alpha$’s, $\alpha \in \Phi$, and $h_\alpha$’s, $\alpha \in \Delta$, with coefficients in $\mathbb{Z}$. It is easy to see that

$$(29.10.14) \quad \tilde{A}(\mathbb{Z}) = \tau(A(\mathbb{Z})), $$

using (29.10.12) and (29.10.13). In particular, the restriction of $\tau$ to $A(\mathbb{Z})$ defines an isomorphism from $A(\mathbb{Z})$ onto $\tilde{A}(\mathbb{Z})$ as Lie algebras over $\mathbb{Z}$, as in Exercise 5 on p150 of [13].

### 29.11 The complex case

Let us continue with the same notation and hypotheses as in Section 29.4, with $k = \mathbb{C}$. As before, we let $\Delta$ be a base for $\Phi$, so that $h_\alpha$, $\alpha \in \Delta$, forms a basis for $A_0$, as a vector space over $\mathbb{C}$. If $\alpha \in \Phi$, then we let $z_\alpha$ be a nonzero element of $A_0$, and ask that $z_\alpha$, $\alpha \in \Phi$, together with $h_\alpha$, $\alpha \in \Delta$, be a Chevalley basis for $A$. Let $\Phi^+ = \Phi^{\Delta^+}$ be the set of positive roots with respect to $\Delta$, which is to say the set of elements of $\Phi$ that can be expressed as linear combinations of elements of $\Delta$ with nonnegative coefficients.

If $\alpha \in \Phi^+$, then put

$$(29.11.1) \quad u_\alpha = z_\alpha - z_{-\alpha}, \quad v_\alpha = i(z_\alpha + z_{-\alpha}).$$
CHAPTER 29. CHEVALLY'S NORMALIZATION

This corresponds to Remark (3) on p52 of [23], and Exercise 7 on p151 of [13].

It is easy to see that

\[(29.11.2) \quad \{u_\alpha, v_\alpha : \alpha \in \Phi^+\} \cup \{ih_\alpha : \alpha \in \Delta\}\]

is a basis for \(A\), as a vector space over \(C\). In particular, the elements of (29.11.2) are linearly independent in \(A\), as a vector space over \(R\). Let \(B\) be the linear span of (29.11.2) in \(A\), as a vector space over \(R\).

Let \(iB\) be the set of elements of \(A\) that can be expressed as \(i\) times an element of \(B\), as usual. Observe that

\[(29.11.3) \quad B + (iB) = A,\]

because (29.11.2) is a basis for \(A\) as a vector space over \(C\). Similarly,

\[(29.11.4) \quad B \cap (iB) = \{0\}.\]

This means that \(A\) may be considered as the complexification of \(B\), as a vector space over \(R\).

Let \(\alpha \in \Delta\) and \(\beta \in \Phi^+\) be given, and observe that

\[(29.11.5) \quad [i h_\alpha, u_\beta]_A = i[h_\alpha, z_\beta]_A - i[h_\alpha, z_{-\beta}]_A = i\lambda_\alpha(\beta) z_\beta - i\lambda_\alpha(-\beta) z_{-\beta} = \lambda_\alpha(\beta) v_\beta.\]

Similarly,

\[(29.11.6) \quad [i h_\alpha, v_\beta]_A = -[h_\alpha, z_\beta]_A - [h_\alpha, z_{-\beta}]_A = -\lambda_\alpha(\beta) z_\beta - \lambda_\alpha(-\beta) z_{-\beta} = -\lambda_\alpha(\beta) u_\beta.\]

Note that the right sides of (29.11.5) and (29.11.6) are elements of \(B\).

If \(\alpha \in \Phi^+\), then

\[(29.11.7) \quad [u_\alpha, v_\alpha]_A = 2i[z_\alpha, z_{-\alpha}]_A = 2i h_\alpha.\]

This is an element of \(B\), because \(h_\alpha\) can be expressed as a linear combination of \(h_\beta, \beta \in \Delta\), with integer coefficients, as in Section 29.4.

If \(\alpha, \beta \in \Phi\) and \(\alpha + \beta \in \Phi\), then there is a unique \(c_{\alpha, \beta} \in C\) such that \([z_\alpha, z_\beta]_A = c_{\alpha, \beta} z_{\alpha + \beta}\), as before. More precisely, \(c_{\alpha, \beta} \in R\), by (29.9.14).

Let \(\alpha, \beta \in \Phi^+\) be given, with \(\alpha \neq \beta\). Note that \(\alpha + \beta \neq 0\) in this case. If \(\alpha + \beta \notin \Phi\), then \(-\alpha - \beta \notin \Phi\), and

\[(29.11.8) \quad [z_\alpha, z_\beta]_A = [z_{-\alpha}, z_{-\beta}]_A = 0.\]

Similarly, if \(\alpha - \beta \notin \Phi\), then \(\beta - \alpha \notin \Phi\), and

\[(29.11.9) \quad [z_{-\alpha}, z_\beta]_A = [z_\alpha, z_{-\beta}]_A = 0.\]

If \(\alpha + \beta \in \Phi\), then

\[(29.11.10) \quad [z_\alpha, z_\beta]_A + [z_{-\alpha}, z_{-\beta}]_A = c_{\alpha, \beta} z_{\alpha + \beta} + c_{-\alpha, -\beta} z_{-\alpha - \beta}\]

\[= c_{\alpha, \beta} z_{\alpha + \beta} - c_{\alpha, \beta} z_{-\alpha - \beta} = c_{\alpha, \beta} u_{\alpha + \beta},\]
29.12. THE CORRESPONDING KILLING FORMS

using (29.8.7) in the second step. Similarly, if \( \alpha - \beta \in \Phi \), then

\[
[z_{-\alpha}, z_{-\beta}]_A + [z_{\alpha}, z_{-\beta}]_A = c_{-\alpha, \beta} z_{-\alpha + \beta} + c_{\alpha, -\beta} z_{\alpha - \beta} = c_{\alpha, -\beta} (z_{\alpha - \beta} - z_{\beta - \alpha}).
\]

This is the same as \( c_{\alpha, -\beta} u_{\alpha - \beta} \) when \( \alpha - \beta \in \Phi^+ \), and \( -c_{\alpha, -\beta} u_{\beta - \alpha} \) when \( \beta - \alpha \in \Phi^+ \).

If \( \alpha + \beta \in \Phi \), then we also have that

\[
i[z_{\alpha}, z_{\beta}]_A - i[z_{-\alpha}, z_{-\beta}]_A = i c_{\alpha, \beta} z_{\alpha + \beta} - i c_{-\alpha, -\beta} z_{-\alpha - \beta} = i c_{\alpha, \beta} (z_{\alpha + \beta} + z_{\beta - \alpha}).
\]

This is the same as \( -c_{\alpha, -\beta} v_{\alpha - \beta} \) when \( \alpha - \beta \in \Phi^+ \), and \( -c_{\alpha, -\beta} v_{\beta - \alpha} \) when \( \beta - \alpha \in \Phi^+ \).

Using these remarks, one can check that

\[
[u_{\alpha}, v_{\beta}]_A, [u_{\alpha}, v_{\beta}]_A, [v_{\alpha}, v_{\beta}]_A \in A
\]

for every \( \alpha, \beta \in \Phi^+ \). It follows that \( B \) is a Lie subalgebra of \( A \), as a Lie algebra over \( \mathbb{R} \), as in Remark (3) on p52 of [23], and Exercise 7 on p151 if [13]. Thus \( A \) may be considered as the complexification of \( B \) as a Lie algebra over \( \mathbb{R} \).

29.12 The corresponding Killing forms

Let us continue with the same notation and hypotheses as in the previous section. If \( x \in A \), then \( \text{ad}_{A,x} \) defines a linear mapping from \( A \) into itself, as a vector space over \( \mathbb{C} \), as in Section 2.4. Similarly, if \( x \in B \), then \( \text{ad}_{B,x} \) defines a linear mapping from \( B \) into itself, as a vector space over \( \mathbb{R} \). Of course, \( \text{ad}_{B,x} \) is the same as the restriction of \( \text{ad}_{A,x} \) to \( B \) in this case.

Remember that the Killing form on \( A \) is defined by

\[
b_A(x, y) = \text{tr}_A(\text{ad}_{A,x} \circ \text{ad}_{A,y})
\]

for every \( x, y \in A \), where the right side is the trace of \( \text{ad}_{A,x} \circ \text{ad}_{A,y} \) as a linear mapping from \( A \) into itself, as a complex vector space. Similarly, the Killing form on \( B \) is defined by

\[
b_B(x, y) = \text{tr}_B(\text{ad}_{B,x} \circ \text{ad}_{B,y})
\]

for every \( x, y \in B \), where the right side is the trace of \( \text{ad}_{B,x} \circ \text{ad}_{B,y} \) as a linear mapping from \( B \) into itself, as a real vector space. If \( x, y \in B \), then

\[
b_A(x, y) = b_B(x, y)
\]
as in Exercise 7 on p151 of [13]. This uses the fact that a basis for $B$ as a real vector space is also a basis for $A$ as a complex vector space, as in the previous section.

If $\alpha, \beta \in \Phi$ and $\alpha + \beta \neq 0$, then

$$(29.12.4) \quad b_A(z_\alpha, z_\beta) = 0,$$

as in Section 17.3. Note that the Lie subalgebra of $A$ that is denoted $B$ in Section 17.3 may be taken to be $A_0$ here. Similarly, if $\alpha \in \Phi$ and $w \in A_0$, then

$$(29.12.5) \quad b_A(z_\alpha, w) = 0,$$

as before.

If $\alpha, \beta \in \Phi^+$ and $\alpha \neq \beta$, then it is easy to see that

$$(29.12.6) \quad b_A(u_\alpha, u_\beta) = b_A(u_\alpha, v_\beta) = b_A(v_\alpha, v_\beta) = 0,$$

using (29.12.4). This implies that

$$(29.12.7) \quad b_B(u_\alpha, u_\beta) = b_B(u_\alpha, v_\beta) = b_B(v_\alpha, v_\beta) = 0,$$

by (29.12.3). Similarly, if $\alpha \in \Phi^+$ and $\beta \in \Delta$, then

$$(29.12.8) \quad b_A(u_\alpha, i h_\beta) = b_A(v_\alpha, i h_\beta) = 0,$$

by (29.12.5). This implies that

$$(29.12.9) \quad b_B(u_\alpha, i h_\beta) = b_B(v_\alpha, i h_\beta) = 0,$$

by (29.12.3).

If $\alpha \in \Phi^+$, then

$$(29.12.10) \quad b_A(u_\alpha, v_\alpha) = i b_A(z_\alpha - z_{-\alpha}, z_\alpha + z_{-\alpha}),$$

by the definition (29.11.1) of $u_\alpha$, $v_\alpha$. Of course,

$$(29.12.11) \quad b_A(z_\alpha, z_\alpha) = b_A(z_{-\alpha}, z_{-\alpha}) = 0,$$

by (29.12.4). This implies that

$$(29.12.12) \quad b_A(u_\alpha, v_\alpha) = 0,$$

because $b_A(\cdot, \cdot)$ is symmetric on $A$. It follows that

$$(29.12.13) \quad b_B(u_\alpha, v_\alpha) = 0,$$

by (29.12.3).

Remember that for each $\alpha \in \Phi$ there is a linear functional $\phi_\alpha$ on $A_0$ such that $\text{ad}_{A_0}(z_\alpha) = \phi_\alpha(w)$ for every $w \in A_0$. We also have that $\phi_\alpha(h_\beta) = \lambda_\beta(\alpha) \in \mathbb{R}$ for every $\beta \in \Phi$, as in (29.4.3). If $w \in A_0 \cap B$, then $w$ can be expressed as a
29.12. **THE CORRESPONDING KILLING FORMS**

linear combination of $i h_{\beta}$, $\beta \in \Delta$, with coefficients in $\mathbf{R}$, by construction. This implies that

(29.12.14) \[ \phi_\alpha(w) \in i \mathbf{R} \]

for every $\alpha \in \Phi$ and $w \in A_0 \cap B$.

If $w, w' \in A_0$, then

(29.12.15) \[ b_A(w, w') = \sum_{\alpha \in \Phi} \phi_\alpha(w) \phi_\alpha(w') \]

This follows from the fact that the $z_\alpha$'s, $\alpha \in \Phi$, together with the $h_{\beta}$'s, $\beta \in \Delta$, form a basis for $A$ as a vector space over $\mathbf{C}$. If $w, w' \in A_0 \cap B$, then

(29.12.16) \[ b_B(w, w') = \sum_{\alpha \in \Phi} \phi_\alpha(w) \phi_\alpha(w') \]

by (29.12.3). In particular,

(29.12.17) \[ b_B(w, w) = \sum_{\alpha \in \Phi} \phi_\alpha(w)^2 \]

for every $w \in A_0 \cap B$.

If $w \in A_0 \cap B$, then every term in the sum on the right side of (29.12.17) is less than or equal to 0, by (29.12.14). If each of these terms is equal to 0, then $w = 0$, as in Section 29.4. This implies that $b_B(\cdot, \cdot)$ is negative-definite on $A_0 \cap B$.

Let $\alpha \in \Phi^+$ be given, and observe that

(29.12.18) \[ b_B([u_\alpha, v_\alpha]_A, i h_\alpha) = 2 b_B(i h_\alpha, i h_\alpha), \]

by (29.11.7). Remember that $h_\alpha \neq 0$, as in Section 29.4, and that $i h_\alpha \in B$, as in the previous section. Thus

(29.12.19) \[ b_B([u_\alpha, v_\alpha]_A, i h_\alpha) < 0, \]

as in the preceding paragraph.

Of course,

(29.12.20) \[ b_B([u_\alpha, v_\alpha]_A, i h_\alpha) = -b_B(v_\alpha, [u_\alpha, i h_\alpha]_A), \]

as in Section 7.9. We also have that

(29.12.21) \[ b_B(v_\alpha, [u_\alpha, i h_\alpha]_A) = -\lambda_\alpha(\alpha) b_B(v_\alpha, v_\alpha) = -2 b_B(v_\alpha, v_\alpha), \]

by (29.11.5). This implies that

(29.12.22) \[ b_B(v_\alpha, v_\alpha) < 0, \]

by (29.12.19).

Similarly,

(29.12.23) \[ b_B([u_\alpha, v_\alpha]_A, i h_\alpha) = b_B([v_\alpha, u_\alpha]_A, i h_\alpha) = b_B(u_\alpha, [v_\alpha, i h_\alpha]_A), \]
as in Section 7.9. Note that
\[(29.12.24) \quad b_B(u_\alpha, [v_\alpha, i h_\alpha], A) = \lambda_\alpha(\alpha) b_B(u_\alpha, u_\alpha) = 2 b_B(u_\alpha, u_\alpha),\]
by (29.11.6). This means that
\[(29.12.25) \quad b_B(u_\alpha, u_\alpha) < 0,\]
because of (29.12.19). One can check more directly that $b_B(u_\alpha, u_\alpha) = b_B(v_\alpha, v_\alpha)$ too.

It follows that $b_B(\cdot, \cdot)$ is negative-definite on $B$, as in Remark (3) on p52 of [23], and Exercise 7 on p151 of [13].

### 29.13 Using other commutative rings

Let us return to the same notation and hypotheses as in Section 29.4. Thus $\Delta$ is a base for $\Phi$, and we let $z_\alpha$ be a nonzero element of $A_\alpha$ for every $\alpha \in \Phi$. Suppose that $z_\alpha, \alpha \in \Phi$, together with $h_\alpha, \alpha \in \Delta$, is a Chevalley basis for $A$, as in Section 29.8. Let $A(Z)$ be the collection of elements of $A$ that can be expressed as a linear combination of the $z_\alpha$’s, $\alpha \in \Phi$, and $h_\alpha$’s, $\alpha \in \Delta$, with integer coefficients. This is a Lie subalgebra of $A$, as a Lie algebra over $Z$, as in Section 29.9.

Let $k_1$ be a commutative ring with a multiplicative identity element. Let us define $A(k_1)$ initially as a free module over $k_1$ with basis consisting of $z_\alpha, \alpha \in \Phi$, and $h_\alpha, \alpha \in \Delta$. These basis elements may also be denoted $z_\alpha k_1, \alpha \in \Phi$, and $h_\alpha k_1, \alpha \in \Delta$, to be more precise. One may consider $A(k_1)$ to be the same as the tensor product of $A(Z)$ and $k_1$, as modules over $Z$. One can use multiplication on $k_1$ to define scalar multiplication by elements of $k_1$ on this tensor product over $Z$, to get a module over $k_1$.

If $n$ is a positive integer, then we can consider the sum of $n$ 1’s in $k_1$, where 1 is the multiplicative identity element in $k_1$. This leads to a natural ring homomorphism from $Z$ into $k_1$, as usual. Using this, we get a mapping from $A(Z)$ into $A(k_1)$ that is linear over $Z$.

It is easy to define the Lie bracket $[\cdot, \cdot]_{A(k_1)}$ initially as a mapping from $A(k_1) \times A(k_1)$ into $A(k_1)$ that is bilinear over $k_1$. The brackets of basis elements can be defined using the corresponding brackets in $A(Z)$, and the natural mapping from $A(Z)$ into $A(k_1)$ described in the preceding paragraph. This can be extended in a unique way to a mapping from $A(k_1) \times A(k_1)$ into $A(k_1)$ that is bilinear over $k$.

Clearly $[\cdot, \cdot]_{A(k_1)}$ is antisymmetric on $A(k_1)$, because $[\cdot, \cdot]_A$ is antisymmetric on $A$. One can check that $[a, a]_{A(k_1)} = 0$ for every $a \in A(k_1)$, using the antisymmetry of $[\cdot, \cdot]_{A(k_1)}$ on $A(k_1)$, and the fact that this property holds when $a$ is a basis element, because of the corresponding property of $[\cdot, \cdot]_A$ on $A$. Similarly, $[\cdot, \cdot]_{A(k_1)}$ satisfies the Jacobi identity on $A(k_1)$, because $[\cdot, \cdot]_A$ satisfies the Jacobi identity on $A$. Thus $A(k_1)$ is a Lie algebra over $k_1$ with respect to $[\cdot, \cdot]_{A(k_1)}$. This may be called a Chevalley algebra, as on p149 of [13]. This is also related to Remark (1) on p51 of [23].
29.14. SOME MAPPINGS ON $A(Z)$

Of course, $A(k_1)$ may be considered as a Lie algebra over $Z$ as well. The natural mapping from $A(Z)$ into $A(k_1)$ mentioned earlier is a homomorphism as a mapping between Lie algebras over $Z$.

Remember that $A(Z)$ does not depend on the base $\Delta$ for $\Phi$, as in Section 29.9. This used the fact that for each $\gamma \in \Phi$, $h_\alpha$ can be expressed as a linear combination of $h_\beta$, with $\beta$ in any other base for $\Phi$, and with integer coefficients. One can use this to check that $A(k_1)$ does not depend on $\Delta$, up to isomorphism, as a Lie algebra over $k_1$. Another base for $\Phi$ essentially only leads to another basis for $A(k_1)$ as a free module over $k_1$. If $A(k_1)$ is considered as the tensor product of $A(Z)$ and $k_1$ over $Z$, then this clearly does not depend on $\Delta$, because $A(Z)$ does not depend on $\Delta$.

Let $e_z$ be another nonzero element of $A$, $\alpha \in \Phi$, for every $\alpha \in \Phi$. Suppose that $e_z$, $\alpha \in \Phi$, together with $h_\alpha$, $\alpha \in \Delta$, is a Chevalley basis for $A$, as in Section 29.8, and let $A(Z)$ be the collection of elements of $A$ that can be expressed as linear combinations of the $e_z$'s, $\alpha \in \Phi$, and $h_\alpha$'s, $\alpha \in \Delta$, then $A(Z) = \tau(A(Z))$, as before. Let $A(k_1)$ be the Lie algebra over $k_1$ defined using the $e_z$'s in the same way as before. One can check that $\tau$ leads to an isomorphism from $A(k_1)$ onto $\hat{A}(k_1)$, as Lie algebras over $k_1$. If $A(k_1)$, $\hat{A}(k_1)$ are considered as tensor products of $A(Z)$, $\hat{A}(Z)$, respectively, and $k_1$ over $Z$, then this follows from the fact that the restriction of $\tau$ to $A(Z)$ defines an isomorphism from $A(Z)$ onto $\hat{A}(Z)$, as Lie algebras over $Z$. This corresponds to a remark on p149 of [13].

29.14 Some mappings on $A(Z)$

Let us continue with the same notation and hypotheses as in Section 29.4 again. Remember that $\Delta$ is a base for $\Phi$, and let $z_\alpha$ be a nonzero element of $A_\alpha$ for every $\alpha \in \Phi$. Suppose that $z_\alpha$, $\alpha \in \Phi$, together with $h_\alpha$, $\alpha \in \Delta$, is a Chevalley basis for $A$, as in Section 29.8, and let $A(Z)$ be the collection of elements of $A$ that can be expressed as a linear combination of the $z_\alpha$'s, $\alpha \in \Phi$, and $h_\alpha$'s, $\alpha \in \Delta$, with integer coefficients. This is a Lie subalgebra of $A$, as a Lie algebra over $Z$, as in Section 29.9.

Let $\alpha \in \Phi$ and a positive integer $m$ be given. We would like to check that $(1/m!)(\text{ad}_{z_\alpha})^m$ maps $A(Z)$ into itself, as in the proposition on p149 of [13]. It suffices to show that

(29.14.1) \[(1/m!)(\text{ad}_{z_\alpha})^m(h_\beta) \in A(Z) \]

for every $\beta \in \Delta$, and that

(29.14.2) \[(1/m!)(\text{ad}_{z_\alpha})^m(z_\beta) \in A(Z) \]

for every $\beta \in \Phi$. 


If \( \beta \in \Delta \), then
\[
\text{ad}_{z_\beta}(h_\beta) = [z_\alpha, h_\beta]_A = -\lambda_\beta(\alpha) z_\alpha,
\]
which is an element of \( A(Z) \). It follows that (ad\(_{z_\alpha}\))\(^m\)(h_\beta) = 0 when \( m \geq 2 \), so that (29.14.1) holds for every \( m \geq 1 \).

Note that (ad\(_{z_\alpha}\))\(^m\)(z_\alpha) = 0 for every \( m \geq 1 \), so that (29.14.2) holds trivially when \( \beta = \alpha \). Remember that ad\(_{z_\alpha}\)(z_\alpha) = \([z_\alpha, z_\alpha]_A = h_\alpha\), as in Section 29.8.

This implies that (29.14.2) holds when \( m = 1 \) and \( \beta = -\alpha \), because \( h_\alpha \in A(Z) \), as in Section 29.9. It follows that
\[
(1^2) (\text{ad}_{z_\alpha})^2(z_\alpha) = (1/2)[z_\alpha, h_\alpha]_A = -(\lambda_\alpha(\alpha)/2) z_\alpha = -z_\alpha,
\]
so that (29.14.2) holds when \( m = 2 \) and \( \beta = -\alpha \). If \( m \geq 3 \), then we get that (ad\(_{z_\alpha}\))\(^m\)(z_\alpha) = 0, so that (29.14.2) holds when \( \beta = -\alpha \).

Suppose now that \( \beta \in \Phi, \beta \neq \pm \alpha \). Thus \( \alpha, \beta \) are linearly independent in \( V \), because \( \Phi \) is reduced, by hypothesis. If \( \gamma \in \Phi \) is not proportional to \( \alpha \), then let \( r(\gamma), q(\gamma) \) be the largest nonnegative integers such that
\[
\gamma = r(\gamma) \alpha + q(\gamma) \alpha \in \Phi,
\]
as in Section 20.5. Remember that \( \gamma + j \alpha \in \Phi \) when \( j \) is an integer that satisfies
\[-r(\gamma) \leq j \leq q(\gamma),
\]
as before. If \( \alpha + \gamma \in \Phi \), then
\[
[z_\alpha, z_\gamma]_A = c_{\alpha, \gamma} z_{\alpha + \gamma} = \pm (r(\gamma) + 1) z_{\alpha + \gamma},
\]
as in Sections 29.8 and 29.9.

Let \( j \) be a nonnegative integer less than or equal to \( q(\beta) \), and observe that
\[
r(\beta + j \alpha) = r(\beta) + j.
\]
If \( 0 \leq j < q(\beta) \), then \( \beta + j \alpha \) and \( \beta + (j + 1) \alpha \) are elements of \( \Phi \), and
\[
[z_\alpha, z_{\beta + j \alpha}]_A = \pm (r(\beta + j \alpha) + 1) z_{\beta + (j + 1) \alpha} = \pm (r(\beta) + j + 1) z_{\beta + (j + 1) \alpha}.
\]
If \( m \) is a positive integer less than or equal to \( q(\beta) \), then we get that
\[
(\text{ad}_{z_\alpha})^m(z_\beta) = \pm \left( \prod_{j=0}^{m-1} (r(\beta) + j + 1) \right) z_{\beta + m \alpha}.
\]
This means that
\[
(1/m!) (\text{ad}_{z_\alpha})^m(z_\beta) = \pm \left( \frac{r(\beta) + m}{m} \right) z_{\beta + m \alpha},
\]
where \( (r(\beta) + m)/m \) is the usual binomial coefficient. It follows that (29.14.2) holds in this case, because the binomial coefficient is an integer.
29.15. SOME RELATED MAPPINGS ON $A(K_1)$

Note that $(\text{ad}_{z_{\beta}})_{q(\beta)+1}(z_{\beta}) = 0$, because $\beta + (q(\beta) + 1) \alpha \not\in \Phi$. This implies that $(\text{ad}_{z_{\beta}})^m(z_{\beta}) = 0$ when $m \geq q(\beta) + 1$, so that (29.14.2) holds for every $m \geq 1$, as desired.

If $c \in \mathbb{Z}$, then

$$\exp(c \text{ad}_{z_{\beta}}) = \sum_{m=0}^{\infty} \left(\frac{c^m}{m!}\right) (\text{ad}_{z_{\beta}})^m$$

(29.14.11)

defines a Lie algebra automorphism of $A$, as in Section 14.11, because $\text{ad}_{z_{\beta}}$ is nilpotent on $A$. We also have that

$$\exp(c \text{ad}_{z_{\beta}})(A(\mathbb{Z})) \subseteq A(\mathbb{Z}),$$

because $(1/m!)(\text{ad}_{z_{\beta}})^m$ maps $A(\mathbb{Z})$ into itself for every $m$, as before. This implies that

$$\exp(c \text{ad}_{z_{\beta}})(A(\mathbb{Z})) = A(\mathbb{Z}),$$

(29.14.12)

by applying the previous statement to $-c$. This corresponds to some remarks on p150 of [13], and is related to Remark (1) on p51 of [23].

29.15 Some related mappings on $A(k_1)$

We continue with the same notation and hypotheses as in the previous section. Let $\alpha \in \Phi$ be given again, and for each positive integer $m$, put

$$Z_{\alpha,m} = (1/m!)(\text{ad}_{z_{\beta}})^m.$$

This defines a linear mapping from $A$ into itself, as a vector space over $k$, which may be interpreted as the identity mapping on $A$ when $m = 0$. Remember that

$$Z_{\alpha,m}(A(\mathbb{Z})) \subseteq A(\mathbb{Z})$$

(29.15.2)

for every $m \geq 0$, as before.

If $l, m$ are nonnegative integers, then

$$\binom{l + m}{l} Z_{\alpha,l+m} = Z_{\alpha,l} \circ Z_{\alpha,m},$$

(29.15.3)

by construction. If $n$ is a nonnegative integer and $a, b \in A$, then

$$\text{ad}_{z_{\beta}}^n([a, b]_A) = \sum_{j=0}^{n} \binom{n}{j} [(\text{ad}_{z_{\beta}})^j(a), (\text{ad}_{z_{\beta}})^{n-j}(b)]_A,$$

(29.15.4)

because $\text{ad}_{z_{\beta}}$ is a derivation on $A$, as in Section 14.10. Equivalently, this means that

$$Z_{\alpha,n}([a, b]_A) = \sum_{j=0}^{n} [Z_{\alpha,j}(a), Z_{\alpha,n-j}(b)]_A.$$

(29.15.5)

Let $k_1$ be a commutative ring with a multiplicative identity element, and let $A(k_1)$ be as in Section 29.13. Note that every mapping from $A(\mathbb{Z})$ into itself
that is linear over \( \mathbb{Z} \) leads to a linear mapping from \( A(k_1) \) into itself that is linear over \( k_1 \). This can be seen in terms of the bases for \( A(\mathbb{Z}) \), \( A(k_1) \) as free modules over \( \mathbb{Z} \), \( k_1 \), respectively, described earlier, or by considering \( A(k_1) \) as the tensor product of \( A(\mathbb{Z}) \) and \( k_1 \), as modules over \( \mathbb{Z} \). This correspondence between mappings on \( A(\mathbb{Z}) \) and on \( A(k_1) \) is linear over \( \mathbb{Z} \), and preserves compositions.

If \( m \) is a nonnegative integer, then let \( Z_{a,m,k_1} \) be the mapping on \( A(k_1) \) that corresponds to the restriction of \( Z_{a,m} \) to \( A(\mathbb{Z}) \) in this way. If \( l \) is another nonnegative integer, then

\[
(l + m) Z_{a,l+m,k_1} = Z_{a,l,k_1} \circ Z_{a,m,k_1},
\]

by (29.15.3). Similarly, if \( n \) is a nonnegative integer and \( a,b \in A(k_1) \), then

\[
Z_{a,n,k_1} ([a,b]_{A(k_1)}) = \sum_{j=0}^{n} [Z_{a,j,k_1}(a), Z_{a,n-j,k_1}(b)]_{A(k_1)},
\]

by (29.15.5). More precisely, it suffices to verify this when \( a, b \) are basis elements of \( A(k_1) \), which can be obtained from the previous case.

If \( t \in k_1 \), then put

\[
E_{a,k_1}(t) = \sum_{l=0}^{\infty} t^l Z_{a,l,k_1}.
\]

Of course, \( Z_{a,l} = 0 \) for all but finitely many \( l \geq 0 \), because \( z_a \) is ad-nilpotent in \( A \). This implies that \( Z_{a,l,k_1} = 0 \) for the same \( l \geq 0 \), so that the right side of (29.15.8) reduces to a finite sum. Thus (29.15.8) defines \( E_{a,k_1}(t) \) as a linear mapping from \( A(k_1) \) into itself.

If \( r \in k_1 \) too, then

\[
E_{a,k_1}(r + t) = \sum_{n=0}^{\infty} (r + t)^n Z_{a,n,k_1} = \sum_{n=0}^{\infty} \sum_{l=0}^{n} \binom{n}{l} t^{n-l} Z_{a,n-l,k_1},
\]

by the binomial theorem. This implies that

\[
E_{a,k_1}(r + t) = \sum_{n=0}^{\infty} \sum_{l=0}^{n} r^l t^{n-l} Z_{a,l,k_1} \circ Z_{a,n-l,k_1},
\]

by (29.15.6). It follows that

\[
E_{a,k_1}(r + t) = E_{a,k_1}(r) \circ E_{a,k_1}(t),
\]

because the right side of (29.15.10) corresponds to the Cauchy product of the series that defined \( E_{a,k_1}(r) \) and \( E_{a,k_1}(t) \). In particular, \( E_{a,k_1}(t) \) is invertible on \( A(k_1) \), with

\[
E_{a,k_1}(t)^{-1} = E_{a,k_1}(-t),
\]

because \( E_{a,k_1}(0) \) is the identity mapping on \( A(k_1) \).
If \( a, b \in A(k_1) \), then

\[
(E_{\alpha, k_1}(t))(\{a, b\}_{A(k_1)}) = \sum_{n=0}^{\infty} t^n Z_{\alpha, n, k_1}([a, b]_{A(k_1)})
\]

(29.15.13)

\[
= \sum_{n=0}^{\infty} t^n \sum_{l=0}^{n} [Z_{\alpha, l, k_1}(a), Z_{\alpha, n-l, k_1}]_{A(k_1)},
\]

using (29.15.7) in the second step. This implies that

\[
(E_{\alpha, k_1}(t))(\{a, b\}_{A(k_1)}) = \sum_{n=0}^{\infty} \sum_{l=0}^{n} [t^l Z_{\alpha, l, k_1}(a), t^{n-l} Z_{\alpha, n-l, k_1}(b)]_{A(k_1)}
\]

(29.15.14)

\[
= [(E_{\alpha, k_1}(t))(a), (E_{\alpha, k_1}(t))(b)]_{A(k_1)},
\]

because the right side is equal to the Cauchy product with respect to \([\cdot, \cdot]_{A(k_1)}\) of the series that define \((E_{\alpha, k_1}(t))(a)\) and \((E_{\alpha, k_1}(t))(b)\). Thus \(E_{\alpha, k_1}(t)\) is an automorphism of \(A(k_1)\), as a Lie algebra over \(k_1\).

This corresponds to some remarks on p150 of [13], and to Exercise 11 on p151 of [13]. This is also related to Remark (1) on p51 of [23].
Chapter 30

Roots and abstract weights

30.1 Abstract weights

Let $V$ be a vector space over the real numbers of positive finite dimension, and let $\Phi$ be a root system in $V$. If $\alpha \in \Phi$, then let $\sigma_\alpha$ be the unique symmetry on $V$ with vector $\alpha$ that maps $\Phi$ onto itself, as usual. Thus $\sigma_\alpha$ can be expressed as $\sigma_\alpha(v) = v - \lambda_\alpha(v)\alpha$, where $\lambda_\alpha$ is a linear functional on $V$ with $\lambda_\alpha(\alpha) = 2$.

As before, we let $\Theta = \Theta_{\Phi}$ be the subgroup of $V$, as a commutative group with respect to addition, generated by $\Phi$. Equivalently, this is the subgroup of $V$ generated by any base for $\Phi$. This may be called the root lattice in $V$ associated to $\Phi$, as on p67 of [13].

Put

(30.1.1) $\Upsilon = \Upsilon_{\Phi} = \{v \in V : \lambda_\alpha(v) \in \mathbb{Z} \text{ for every } \alpha \in \Phi\}.$

The elements of $\Upsilon$ may be called (abstract) weights with respect to $\Phi$, as on p67 of [13]. This corresponds to $P$ on p63 of [23], where the root system is associated to a finite-dimensional semisimple Lie algebra. Note that $\Upsilon$ is a subgroup of $V$, as a commutative group with respect to addition. We also have that

(30.1.2) $\Theta \subseteq \Upsilon,$

by the definition of a root system.

Let $V'$ be the dual space of linear functionals on $V$, and remember that

(30.1.3) $\Phi' = \{\lambda_\alpha : \alpha \in \Phi\}$

is a root system in $V'$, as in Section 19.8. If $\alpha \in \Phi$ and $2\alpha \in \Phi$, then $\sigma_{2\alpha} = \sigma_\alpha$, and $\lambda_{2\alpha} = \lambda_\alpha/2$, as in Section 19.13. Let $\Delta$ be a base for $\Phi$, and put

(30.1.4) $\Delta_1 = \{\beta \in \Delta : 2\beta \not\in \Phi\}, \quad \Delta_2 = \{\beta \in \Delta : 2\beta \in \Phi\}.$

Note that $\Delta_2 = \emptyset$ when $\Phi$ is reduced, and put

(30.1.5) $\Delta'_1 = \{\lambda_\beta : \beta \in \Delta_1\}, \quad \Delta'^*_2 = \{\lambda_\beta/2 : \beta \in \Delta_2\}.$
30.2. SOME REMARKS ABOUT LATTICES

Remember that $\Delta'_1 \cup \Delta'_2$ is a base for $\Phi'$, as in Section 19.13.

It follows that

$$\Upsilon = \{v \in V : \lambda_\beta(v) \in \mathbb{Z} \text{ for every } \beta \in \Delta_1, \text{ and } \lambda_\beta(v)/2 \in \mathbb{Z} \text{ for every } \beta \in \Delta_2\}.$$  

Equivalently,

$$\Upsilon = \{v \in V : \lambda(v) \in \mathbb{Z} \text{ for every } \lambda \in \Delta'_1 \cup \Delta'_2\}.$$  

In particular, $\Delta'_1 \cup \Delta'_2$ is a basis for $V'$, as a vector space over $\mathbb{R}$. If $v \in V$, then

$$\lambda \mapsto \lambda(v)$$

defines a real-valued function on $\Delta'_1 \cup \Delta'_2$. The mapping from $v \in V$ to (30.1.8) defines a one-to-one linear mapping from $V$ onto the space of real-valued functions on $\Delta'_1 \cup \Delta'_2$. Thus, for each $\mu \in \Delta'_1 \cup \Delta'_2$, there is a unique $v_\mu \in V$ such that

$$\lambda(v_\mu) = 1 \text{ when } \lambda = \mu$$

$$= 0 \text{ for every } \lambda \in \Delta'_1 \cup \Delta'_2 \text{ with } \lambda \neq \mu.$$  

The $v_\mu$'s, $\mu \in \Delta'_1 \cup \Delta'_2$, form a basis for $V$, which is dual to $\Delta'_1 \cup \Delta'_2$.

Using (30.1.7), we get that

$$v_\lambda \in \Upsilon$$

for every $\lambda \in \Delta'_1 \cup \Delta'_2$. The $v_\lambda$'s, $\lambda \in \Delta'_1 \cup \Delta'_2$, are called the fundamental (dominant) weights with respect to $\Delta$, as on p67 of [13], and Remark (2) on p62 of [23]. Note that $\Upsilon$ consists exactly of linear combinations of the $v_\lambda$'s, $\lambda \in \Delta'_1 \cup \Delta'_2$, with integer coefficients.

30.2 Some remarks about lattices

Let $V$ be a vector space over the real numbers of positive finite dimension $n$, and let $u_1, \ldots, u_n$ and $w_1, \ldots, w_n$ be bases for $V$. Also let $A$ and $B$ be the subgroups of $V$, as a commutative group with respect to addition, generated by the $u_j$'s and $w_j$'s, respectively. Equivalently, $A$ and $B$ consist of the elements of $V$ that can be expressed as linear combinations of the $u_j$'s and $w_j$'s, respectively, with integer coefficients.

Suppose that

$$B \subseteq A.$$  

This means that for each $l = 1, \ldots, n$, $w_l$ can be expressed as a linear combination of the $u_j$'s with integer coefficients. Let $T$ be the linear mapping from $V$ onto itself with

$$T(w_l) = w_l.$$
CHAPTER 30. ROOTS AND ABSTRACT WEIGHTS

for every \( l = 1, \ldots, n \). Note that \( T \) is invertible, so that \( \det T \neq 0 \). We also have that

\[
T(A) = B,
\]

by construction.

Of course, \( T \) corresponds to an \( n \times n \) matrix, with respect to the basis \( u_1, \ldots, u_n \) of \( V \). The entries of this matrix are obtained by expressing \( w_l \) as a linear combination of the \( u_j \)'s for each \( l = 1, \ldots, n \). In particular, the entries of the matrix associated to \( T \) are integers in this case. This implies that

\[
\det T \in \mathbb{Z}.
\]

Using Cramer’s rule, we get that \( (\det T) T^{-1} \) corresponds to a matrix with integer entries as well. This means that

\[
(\det T) T^{-1}A \subseteq A.
\]

It follows that

\[
(\det T) A \subseteq T(A) = B.
\]

This implies that the quotient group \( A/B \) has only finitely many elements.

Because \( V \) is isomorphic to \( \mathbb{R}^n \) as a vector space over \( \mathbb{R} \), one can define \( n \)-dimensional volumes on \( V \) as on \( \mathbb{R}^n \). These volumes are determined up to multiplication by a positive real number, depending on the isomorphism with \( \mathbb{R}^n \). These volumes are invariant under translations on \( V \), and are changed by \( T \) by a factor of \( |\det T| \), because of the analogous properties on \( \mathbb{R}^n \).

It is well known that

\[
\#(A/B) = |\det T|,
\]

where the left side is the number of elements of \( A/B \). To see this, one can consider the quotients \( V/A \) and \( V/B \) of \( V \) by \( A \) and \( B \), respectively. There is a natural quotient mapping from \( V/A \) onto \( V/B \), because of (30.2.1). The natural quotient mapping from \( V \) onto \( V/A \) is the same as the composition of the natural quotient mapping from \( V \) onto \( V/B \) with the quotient mapping from \( V/B \) onto \( V/A \). One can define \( n \)-dimensional volumes on \( V/A \) and \( V/B \) that correspond locally to volumes on \( V \) under the quotient mappings. In particular, volumes on \( V/A \) correspond locally to volumes on \( V/B \) under the natural quotient mapping. This implies that the ratio of the volumes of \( V/B \) and \( V/A \) is equal to \( \#(A/B) \).

There is also a natural one-to-one mapping from \( V/A \) onto \( V/B \) induced by \( T \). Because volumes on \( V \) are changed by \( T \) by a factor of \( |\det T| \), volumes on the quotients are changed in the same way by the mapping induced by \( T \), where the volumes on the quotients correspond locally to volumes on \( V \) as before. This means that the ratio of the volumes of \( V/B \) and \( V/A \) is equal to \( |\det T| \), which implies (30.2.7).

Let us return now to the situation considered in the previous section. Note that \( \Theta \) is generated, as a subgroup of \( V \), by \( \Delta \), which is a basis for \( V \). We also have that \( \Upsilon \) is generated, as a subgroup of \( V \), by the \( v_\lambda \)'s, \( \lambda \in \Delta_1 \cup \Delta_2 \), which form a basis for \( V \) too. Remember that \( \Theta \subseteq \Upsilon \), so that the quotient \( \Upsilon/\Theta \) is
30.3. RELATED POSITIVITY CONDITIONS

Let us go back to the notation and hypotheses in Section 30.1. In particular, $\Delta$ is a base for $\Phi$, and we put

$$\Upsilon^+ = \Upsilon^+_{\Phi,\Delta} = \{v \in \Upsilon : \lambda_\alpha(v) \geq 0 \text{ for every } \alpha \in \Delta\}.$$  (30.3.1)

The elements of $\Upsilon^+$ are said to be dominant weights with respect to $\Delta$, as on p67 of [13]. If $v \in \Upsilon$ satisfies

$$\lambda_\alpha(v) > 0 \text{ for every } \alpha \in \Delta,$$  (30.3.2)

then $v$ is said to be strongly dominant.

Equivalently, it is easy to see that

$$\Upsilon^+ = \{v \in \Upsilon : \lambda(v) \geq 0 \text{ for every } \lambda \in \Delta'_1 \cup \Delta'_2\}.$$  (30.3.3)
Similarly, (30.3.2) holds if and only if

\[(v) > 0 \quad \text{for every } \Delta'_1 \cup \Delta'_2.\]  

If \(\mu \in \Delta'_1 \cup \Delta'_2\) and \(v_\mu \in \Upsilon\) is as in (30.1.9), then we get that

\[v_\mu \in \Upsilon^+,\]  

by (30.3.3).

Let \(\Phi^+ = \Phi^{\Delta^+}\) be the set of \(\alpha \in \Phi\) that can be expressed as linear combinations of elements of \(\Delta\) with nonnegative coefficients, as usual. Put

\[(\Phi')^+ = \{\lambda_\alpha : \alpha \in \Phi^+\}.\]  

Remember that \(\Delta'_1 \cup \Delta'_2\) is a base for \(\Phi',\) as a root system in \(V'.\) One can check directly that

\[\Delta'_1 \cup \Delta'_2 \subseteq (\Phi')^+,\]  

because \(\Delta \subseteq \Phi^+\).

In fact, \((\Phi')^+\) consists exactly of the elements of \(\Phi'\) that can be expressed as linear combinations of elements of \(\Delta'_1 \cup \Delta'_2\) with nonnegative coefficients. This is basically implicit in the proof of the fact that \(\Delta'_1 \cup \Delta'_2\) is a base for \(\Phi',\) as in Section 19.13. More precisely, in the proof an inner product on \(V\) that is invariant under the Weyl group of \(\Phi\) was used to identify \(V\) and \(V'.\) Using this identification, \(\Delta'_1 \cup \Delta'_2\) was shown to correspond to the construction of a base for a root system in Section 19.11. The set of positive roots with respect to this base is also determined in this construction, as before. In this case, the set of positive roots corresponds exactly to \((\Phi')^+\). This is because of the way that the linear functional \(\tau\) is defined in the proof in Section 19.13.

If \(v \in \Upsilon^+,\) then it follows that

\[\lambda_\alpha(v) \geq 0 \quad \text{for every } \alpha \in \Phi^+.\]  

This corresponds to the statement following Theorem 3 on p60 of [23]. Similarly, if \(v \in \Upsilon\) is strongly dominant, then

\[\lambda_\alpha(v) > 0 \quad \text{for every } \alpha \in \Phi^+.\]  

Let \((\cdot, \cdot)\) be an inner product on \(V\) that is invariant under the Weyl group of \(\Phi.\) If \(\alpha \in \Phi,\) then \(\sigma_\alpha\) is the reflection on \(V\) associated to \(\alpha\) with respect to \((\cdot, \cdot),\) so that \(\lambda_\alpha(v) = 2(v, \alpha)(\alpha, \alpha)^{-1},\) as before. Under these conditions,

\[\Upsilon^+ = \{v \in \Upsilon : (v, \alpha) \geq 0 \text{ for every } \alpha \in \Delta\}.\]  

Similarly, \(v \in \Upsilon\) is strongly dominant if and only if

\[(v, \alpha) > 0 \quad \text{for every } \alpha \in \Delta.\]  

This corresponds to some remarks about Weyl chambers on p67 of [13].
30.4. AUTOMORPHISMS AND \( \Upsilon \)

If \( v \in V \), then (30.3.8) holds if and only if
\[
(v, \alpha) \geq 0 \quad \text{for every } \alpha \in \Phi^+. 
\]
(30.3.12)

In particular, this holds when \( v \in \Upsilon^+ \), as before. Alternatively, (30.3.12) holds when \( v \in V \) satisfies \((v, \alpha) \geq 0\) for every \( \alpha \in \Delta \), by definition of \( \Phi^+ \).

Similarly, \( v \in V \) satisfies (30.3.9) if and only if
\[
(v, \alpha) > 0 \quad \text{for every } \alpha \in \Phi^+. 
\]
(30.3.13)

This holds when \( v \in \Upsilon \) is strongly dominant, as before. Alternatively, (30.3.11) implies (30.3.13), by definition of \( \Phi^+ \).

30.4 Automorphisms and \( \Upsilon \)

Let us continue with the notation and hypotheses in Section 30.1 again. Remember that a one-to-one linear mapping \( T \) from \( V \) onto itself is an automorphism of \( \Phi \) when \( T(\Phi) = \Phi \), and that \( \text{Aut}(\Phi) \) denotes the group of these automorphisms. If \( T \in \text{Aut}(\Phi) \), then
\[
T(\Theta) = \Theta. 
\]
(30.4.1)

In this case,
\[
\lambda_{T(\alpha)}(T(v)) = \lambda_\alpha(v) 
\]
(30.4.2)

for every \( \alpha \in \Phi \) and \( v \in V \), as in Section 19.5. One can use this to check that
\[
T(\Upsilon) = \Upsilon. 
\]
(30.4.3)

Remember that the quotient group \( \Upsilon/\Theta \) has only finitely many elements, as in Section 30.2. If \( T \in \text{Aut}(\Phi) \), then \( T \) induces an automorphism \( \tilde{T} \) of \( \Upsilon/\Theta \), by (30.4.1) and (30.4.3).

If \( \alpha \in \Phi \) and \( v \in \Upsilon \), then
\[
\sigma_\alpha(v) - v = -\lambda_\alpha(v) \alpha \in \Theta. 
\]
(30.4.4)

Of course, \( \sigma_\alpha \in \text{Aut}(\Phi) \), so that \( \sigma_\alpha \) induces an automorphism \( \tilde{\sigma}_\alpha \) on \( \Upsilon/\Theta \), as in the preceding paragraph. It follows from (30.4.4) that \( \tilde{\sigma}_\alpha \) is the identity mapping on \( \Upsilon/\Theta \). If \( \sigma \) is any element of the Weyl group of \( \Phi \), then \( \sigma \in \text{Aut}(\Phi) \), and we get that \( \tilde{\sigma} \) is the identity mapping on \( \Upsilon/\Theta \). Equivalently, this means that
\[
\sigma(v) - v \in \Theta 
\]
for every \( v \in \Upsilon \).

Let \((\cdot, \cdot)\) be an inner product on \( V \) that is invariant under the Weyl group of \( \Phi \). Remember that \( \Delta \) is a base for \( \Phi \), and suppose for the rest of the section that \( \Phi \) is reduced, as a root system in \( V \). If \( v \in V \), then there is an element \( \sigma \) of the Weyl group of \( \Phi \) such that
\[
(\sigma(v), \alpha) \geq 0 \quad \text{for every } \alpha \in \Delta, 
\]
(30.4.6)
as in Section 20.9. If \( v \in \mathcal{Y} \), then \( \sigma(v) \in \mathcal{Y}^+ \), and it follows that

\[
\sigma(v) \in \mathcal{Y}^+,
\]

as in (30.3.10). This corresponds to part of Lemma A on p68 of [13].

As in Section 21.12, a linear functional \( \mu \) on \( V \) is said to be nonnegative with respect to \( \Delta \) if

\[
\mu(\alpha) \geq 0 \quad \text{for every } \alpha \in \Delta.
\]

If \( v \in V \), then put

\[
\mu_v(w) = (v, w)
\]

for every \( w \in V \), which defines a linear functional on \( V \). If \( \sigma \) is in the Weyl group of \( \Phi \), then

\[
\mu_{\sigma(v)}(w) = (\sigma(v), w) = (v, \sigma^{-1}(w))
\]

for every \( w \in V \), which is to say that

\[
\mu_{\sigma(v)} = \mu \circ \sigma^{-1}.
\]

Suppose that \( \mu_v \) is nonnegative with respect to \( \Delta \), and that \( \mu_{\sigma(v)} \) is nonnegative with respect to \( \Delta \) for some \( \sigma \) in the Weyl group of \( \Phi \). Under these conditions,

\[
\mu_{\sigma(v)} = \mu_v,
\]

as in Section 21.12. Equivalently, this means that

\[
\sigma(v) = v.
\]

In particular, this holds when \( v \in \mathcal{Y}^+ \) and \( \sigma(v) \in \mathcal{Y}^+ \) for some \( \sigma \) in the Weyl group of \( \Phi \). This is another part of Lemma A on p68 of [13].

Suppose again that \( \mu_v \) is nonnegative with respect to \( \Delta \) for some \( v \in V \), and that \( \mu_{\sigma(v)} \) is nonnegative with respect to \( \Delta \) for some \( \sigma \) in the Weyl group of \( \Phi \). As in Section 21.12, this implies that \( \sigma \) can be expressed as the composition of finitely many reflections \( \sigma_\alpha \) with \( \alpha \in \Delta \) and

\[
\mu_v = \mu_v \circ \sigma_\alpha.
\]

Observe that (30.4.14) implies that

\[
\mu_v(\alpha) = -\mu_v(\alpha),
\]

so that \( \mu_v(\alpha) = 0 \). This is not possible when \( \mu_v \) is strictly positive with respect to \( \Delta \), in the sense that

\[
\mu_v(\alpha) = (v, \alpha) > 0 \quad \text{for every } \alpha \in \Delta.
\]

In this case, we get that \( \sigma \) is the identity mapping on \( V \).

If \( v \in \mathcal{Y} \) is strongly dominant and \( \sigma(v) \in \mathcal{Y}^+ \) for some \( \sigma \) in the Weyl group of \( \Phi \), then it follows that \( \sigma \) is the identity mapping on \( V \). This is another part of Lemma A on p68 of [13].
30.5 Some partial orderings on $V$

Let us continue with the usual notation and hypotheses in Section 30.1. Thus $\Delta$ is a base for $\Phi$, and a basis for $V$ in particular. Consider the binary relation $\preceq \Delta$ defined on $V$ by putting $v \preceq w$ when $v, w \in V$ and either $v = w$, or $w - v$ can be expressed as a finite sum of elements of $\Delta$. This defines a partial ordering on $V$, as on p47 of [13].

Similarly, consider the binary relation $\preceq^R \Delta$ defined on $V$ by putting $v \preceq^R w$ when $v, w \in V$ and $w - v$ can be expressed as a linear combination of elements of $\Delta$ with nonnegative coefficients. This defines a partial ordering on $V$ too, as mentioned in Exercise 14 on p55 of [13]. Of course, $v \preceq w$ implies $v \preceq^R w$ (30.5.1).

Let $v, w \in V$ and $\alpha \in \Delta$ be given, and note that

\[(30.5.2)\] 

\[v - \sigma_\alpha(v) = \lambda_\alpha(v) \alpha.\]

In this case, $\sigma_\alpha(v) \preceq^R v$ when $\lambda_\alpha(v) \geq 0$, and $v \preceq^R \sigma_\alpha(v)$ when $\lambda_\alpha(v) \leq 0$. If $v \in \mathcal{Y}$, then we get that $\sigma_\alpha(v) \preceq v$ when $\lambda_\alpha(v) \geq 0$, and $v \preceq \sigma_\alpha(v)$ when $\lambda_\alpha(v) \leq 0$.

Let $v \in V$ be given again, and suppose that $\sigma$ is an element of the Weyl group of $\Phi$ such that $\sigma(v)$ is maximal with respect to $\preceq^R$ among the images of $v$ under elements of the Weyl group. If $\alpha \in \Delta$ and $\sigma(v) \preceq^R \sigma_\alpha(\sigma(v))$, then it follows that $\sigma_\alpha(\sigma(v)) = \sigma(v)$. This means that

\[(30.5.3)\] 

\[\lambda_\alpha(\sigma(v)) \geq 0 \quad \text{for every } \alpha \in \Delta,
\]
as in the preceding paragraph.

Let $(\cdot, \cdot)$ be an inner product on $V$ that is invariant under the Weyl group of $\Phi$. Observe that (30.5.3) is the same as saying that (30.4.6) holds.

Of course, there are only finitely many images of $v$ in $V$ under elements of the Weyl group of $\Phi$, because the Weyl group has only finitely many elements. This implies that there is an element $\sigma$ of the Weyl group such that $\sigma(v)$ is maximal with respect to $\preceq^R$, because every nonempty finite partially-ordered set has a maximal element. This gives another way to get the existence of an element $\sigma$ of the Weyl group such that (30.4.6) holds, as in Exercise 14 on p55 of [13].

Let $\sigma^0$ be any element of the Weyl group of $\Phi$. Consider the set of elements $\sigma^1$ of the Weyl group such that

\[(30.5.4)\] 

\[\sigma^0(v) \preceq^R \sigma^1(v).\]

Let $\sigma$ be an element of this set such that $\sigma(v)$ is maximal with respect to $\preceq^R$ among the images of $v$ under the elements of this set, which exists as before. Thus

\[(30.5.5)\] 

\[\sigma^0(v) \preceq^R \sigma(v),\]
and it is easy to see that $\sigma(v)$ is maximal among all images of $v$ under elements of the Weyl group with respect to $\preceq^R$. This implies that (30.4.6) holds, as before.

Suppose from now on in this section that $\Phi$ is reduced, as a root system in $V$. Suppose also that $v \in V$ satisfies

$$\tag{30.5.6} (v, \alpha) \geq 0 \quad \text{for every } \alpha \in \Delta.$$ 

This is the same as saying that the linear functional $\mu_v$ on $V$ associated to $v$ as in (30.4.9) is nonnegative with respect to $\Delta$. If $\sigma$ is an element of the Weyl group that satisfies (30.4.6), then $\mu_{\sigma(v)}$ is nonnegative with respect to $\Delta$ too. This implies that $\sigma(v) = v$, as before.

If $\sigma^0$ is any element of the Weyl group, then there is an element $\sigma$ of the Weyl group that satisfies (30.4.6) and (30.5.5). Under these conditions, we get that

$$\tag{30.5.7} \sigma^0(v) \preceq^R v,$$

because $\sigma(v) = v$, as in the preceding paragraph.

Suppose that $v$ is a dominant weight, so that $v \in \Upsilon$ and (30.5.6) holds, as in (30.3.10). Note that $v - \sigma^0(v) \in \Theta$, as in (30.4.5). In this case, (30.5.7) implies that

$$\tag{30.5.8} \sigma^0(v) \preceq v.$$ 

This corresponds to another part of Lemma A on p68 of [13].

### 30.6 Some more properties of $\Upsilon$

Let us continue with the usual notation and hypotheses in Section 30.1 again. In particular, $\Delta$ is a base for $\Phi$, so that $\Upsilon^+$ can be defined as in Section 30.3. Let $\preceq$ be defined on $V$ as in the previous section, and let $(\cdot, \cdot)$ be an inner product on $V$ that is invariant under the Weyl group of $\Phi$.

Let $v \in \Upsilon^+$ be given, and suppose that

$$\tag{30.6.1} w \in \Upsilon^+ \text{ satisfies } w \preceq v.$$ 

This implies that $v + w \in \Upsilon^+$, and that $v - w$ can be expressed as a sum of elements of $\Delta$. It follows that

$$\tag{30.6.2} (v + w, v - w) \geq 0,$$

because $(v + w, \alpha) \geq 0$ for every $\alpha \in \Delta$. This means that

$$\tag{30.6.3} (v, v) - (w, w) \geq 0,$$

which is to say that $(w, w) \leq (v, v)$. Using this, one can check that there are only finitely many $w$ as in (30.6.1), which is Lemma B at the top of p70 of [13].
Let $\Phi^+$ be the set of elements of $\Phi$ that can be expressed as linear combinations of elements of $\Delta$ with nonnegative coefficients, as before. Suppose from now on in this section that $\Phi$ is reduced as a root system in $V$, and put

$$\rho = \frac{1}{2} \sum_{\gamma \in \Phi^+} \gamma,$$

as in Section 19.12. If $\alpha \in \Delta$, then

$$\sigma_{\alpha}(\rho) = \rho - \alpha,$$

as before. Equivalently, this means that

$$\lambda_{\alpha}(\rho) = 1$$

for every $\alpha \in \Delta$. This shows that

$$\rho \in \Upsilon^+,$$

and more precisely that $\rho$ is strongly dominant with respect to $\Delta$, as in Lemma A around the middle of p70 of [13].

Because $\Phi$ is reduced, we have that $\Delta_1 = \Delta$ and $\Delta_2 = \emptyset$, in the notation of Section 30.1. If $\beta \in \Delta$, then let $v_\beta = v_{\lambda_\beta} \in \Upsilon$ be as in Sections 30.1 and 30.2. Using (30.2.10) and (30.6.6), we get that

$$\rho = \sum_{\beta \in \Delta} v_\beta.$$

This is another part of Lemma A around the middle of p70 of [13].

Let $u \in \Upsilon^+$ be given, and suppose that $w = \sigma^{-1}(u)$ for some $\sigma$ in the Weyl group of $\Phi$. Under these conditions, Lemma B around the middle of p70 of [13] states that

$$(w + \rho, w + \rho) \leq (u + \rho, u + \rho),$$

with equality only when $u = w$. To see this, observe first that

$$(w + \rho, w + \rho) = (\sigma(w + \rho), \sigma(w + \rho)) = (u + \sigma(\rho), u + \sigma(\rho)).$$

This implies that

$$(w + \rho, w + \rho) = (u + \rho, u + \rho) - 2(u, \rho - \sigma(\rho)),$$

because $(\sigma(\rho), \sigma(\rho)) = (\rho, \rho)$.

Note that

$$\sigma(\rho) \preceq \rho,$$

by (30.5.8) and (30.6.7). Thus $\rho - \sigma(\rho)$ is either equal to 0, or a sum of elements of $\Delta$. It follows that

$$(u, \rho - \sigma(\rho)) \geq 0,$$
because \(u \in \mathbb{T}^+\). This implies (30.6.9), using (30.6.11). We also get that equality holds in (30.6.9) if and only if

\[(30.6.14) \quad (u, \rho - \sigma(\rho)) = 0.\]

Clearly

\[(30.6.15) \quad (u, \rho - \sigma(\rho)) = (u, \rho) - (u, \sigma(\rho)) = (u, \rho) - (\sigma^{-1}(u), \rho) = (u - w, \rho).\]

We also have that

\[(30.6.16) \quad w \preceq u,\]

as in (30.5.8), because \(u \in \mathbb{T}^+\). This means that either \(u = w\), or \(u - w\) is a sum of elements of \(\Delta\). In the second case, we would get that

\[(30.6.17) \quad (u, \rho - \sigma(\rho)) = (u - w, \rho) > 0,\]

because \(\rho\) is strongly dominant with respect to \(\Delta\), as before. Thus (30.6.14) holds only when \(u = w\), as desired.

### 30.7 Saturated subsets of \(\Upsilon\)

Let us continue with the same notation and hypotheses as in Sections 30.1 and 30.3, and let \(\preceq\) be defined on \(V\) as in Section 30.5. A subset \(\Pi\) of \(\Upsilon\) is said to be saturated if for every \(v \in \Pi\) and \(\alpha \in \Phi\) we have that

\[(30.7.1) \quad v - j \alpha \in \Pi\]

for every integer \(j\) between 0 and \(\lambda_\alpha(v)\), as on p70 of [13]. Of course, this holds trivially when \(j = 0\), and it is also supposed to hold when \(j = \lambda_\alpha(v)\).

If \(\alpha \in \Phi\), then it follows that \(\sigma_\alpha(\Pi) \subseteq \Pi\). This implies that \(\sigma_\alpha(\Pi) = \Pi\), because \(\sigma_\alpha\) is its own inverse on \(V\). Thus

\[(30.7.2) \quad \sigma(\Pi) = \Pi\]

for every \(\sigma\) in the Weyl group of \(\Phi\) when \(\Pi\) is saturated, as on p70 of [13].

A saturated set \(\Pi \subseteq \Upsilon\) is said to have highest weight \(v \in \mathbb{T}^+\) if \(v \in \Pi\) and

\[(30.7.3) \quad w \preceq v \quad \text{for every } w \in \Pi,\]

as on p70 of [13]. It is easy to see that \(\Pi = \{0\}\) is saturated, with highest weight 0, as in Example (1) on p70 of [13].

Let us check that \(\Pi = \Phi \cup \{0\}\) is saturated, as in Example (2) on p70 of [13]. If \(v = 0\), then \(\lambda_\alpha(v) = 0\) for every \(\alpha \in \Phi\), and (30.7.1) holds trivially. Otherwise, let \(v, \alpha \in \Phi\) be given, and note that (30.7.1) holds when \(j = \lambda_\alpha(v)\), because \(\sigma_\alpha(\Pi) \subseteq \Phi\), by the definition of a root system. If \(v\) and \(\alpha\) are not proportional
in $V$, then it follows that (30.7.1) holds when $j$ is between 0 and $\lambda_\alpha(v)$, as in Section 20.5. If $v$ and $\alpha$ are proportional in $V$, then this can be verified directly.

Let us say that an element of $\Phi$ is maximal with respect to $\Delta$ if it is maximal with respect to $\preceq$. This is equivalent to the definition of a maximal root in Section 20.8. If $\alpha \in \Phi$, then the set of $\beta \in \Phi$ such that $\alpha \preceq \beta$ is a nonempty finite set. It follows that this set has a maximal element $\beta_0$ with respect to $\preceq$, and it is easy to see that $\beta_0$ is also maximal in $\Phi$.

Suppose for the moment that $\Phi$ is irreducible as a root system in $V$. This implies that $\Phi$ has a unique maximal element $\alpha_0$, as in Section 20.8. If $\alpha \in \Phi$, then

\[(30.7.4) \quad \alpha \preceq \alpha_0,\]

by the remarks in the preceding paragraph, and which was also mentioned in Section 20.8. It is easy to see that (30.7.4) holds with $\alpha = 0$ too.

Let $(\cdot, \cdot)$ be an inner product on $V$ that is invariant under the Weyl group of $\Phi$. If $\gamma \in \Delta$, then

\[(30.7.5) \quad (\alpha_0, \gamma) \geq 0,\]

as in Section 20.8. This means that

\[(30.7.6) \quad \alpha_0 \in \mathbb{Y}^+,\]

as in Section 30.3. It follows that $\Pi = \Phi \cup \{0\}$ has highest weight $\alpha_0$ under these conditions, as on p70 of [13].

Suppose that $\Pi \subseteq \mathbb{Y}$ is saturated, and has highest weight $v \in \mathbb{Y}^+$. Using the remarks near the beginning of the previous section, we get that $\Pi \cap \mathbb{Y}^+$ has only finitely many elements. Suppose also that $\Phi$ is reduced, as a root system in $V$. If $w \in \Pi$, then there is an element $\sigma$ in the Weyl group of $\Phi$ such that $\sigma(w) \in \mathbb{Y}^+$, as in Section 30.4. We also have that $\sigma(w) \in \Pi$, as in (30.7.2). It follows that $\Pi$ has only finitely many elements under these conditions, because $\Pi \cap \mathbb{Y}^+$ and the Weyl group of $\Phi$ have only finitely many elements. This is Lemma A near the bottom of p70 of [13].

### 30.8 More on highest weights

We continue with the same notation and hypotheses as in Sections 30.1 and 30.3, and to let $\preceq$ be defined on $V$ as in Section 30.5. We also ask that $\Phi$ be reduced as a root system in $V$ in this section. Let $(\cdot, \cdot)$ be an inner product on $V$ that is invariant under the Weyl group of $\Phi$.

Suppose that $\Pi \subseteq \mathbb{Y}$ is saturated, with highest weight $v \in \mathbb{Y}^+$. If $u \in \mathbb{Y}^+$ and $u \preceq v$, then Lemma B at the bottom of p70 of [13] states that

\[(30.8.1) \quad u \in \Pi.\]

To see this, we shall consider elements $w$ of $\Pi$ of the form

\[(30.8.2) \quad w = u + \sum_{\alpha \in \Delta} c_\alpha \alpha,\]
where \( c_\alpha \) is a nonnegative integer for each \( \alpha \in \Delta \). Of course, if this happens with \( c_\alpha = 0 \) for each \( \alpha \in \Delta \), then (30.8.1) holds, as desired. Note that \( w = v \) can be expressed as in (30.8.2), by hypothesis.

Suppose that \( w \in \Pi \) is of the form (30.8.2), and that \( c_\alpha > 0 \) for some \( \alpha \in \Delta \). Thus \( \sum_{\alpha \in \Delta} c_\alpha \alpha \neq 0 \), which implies that

\[
(30.8.3) \quad \left( \sum_{\alpha \in \Delta} c_\alpha \alpha, \sum_{\beta \in \Delta} c_\beta \beta \right) > 0.
\]

It follows that there is a \( \beta_0 \in \Delta \) such that \( c_{\beta_0} > 0 \) and

\[
(30.8.4) \quad \left( \sum_{\alpha \in \Delta} c_\alpha \alpha, \beta_0 \right) > 0.
\]

Equivalently,

\[
(30.8.5) \quad \lambda_{\beta_0} \left( \sum_{\alpha \in \Delta} c_\alpha \alpha \right) > 0.
\]

We also have that

\[
(30.8.6) \quad \lambda_{\beta_0}(u) \geq 0,
\]

because \( u \in \Upsilon^+ \), by hypothesis. Combining this with (30.8.5), we obtain that

\[
(30.8.7) \quad \lambda_{\beta_0}(w) > 0.
\]

This means that \( \lambda_{\beta_0}(w) \geq 1 \), because \( w \in \Pi \subseteq \Upsilon \), so that \( \lambda_{\beta_0}(w) \in \mathbb{Z} \). It follows that

\[
(30.8.8) \quad w - \beta \in \Pi,
\]

because \( w \in \Pi \) and \( \Pi \) is saturated.

Note that \( c_{\beta_0} \geq 1 \), because \( c_{\beta_0} > 0 \) and \( c_{\beta_0} \in \mathbb{Z} \). This means that \( w - \beta \) can be expressed as in (30.8.2) as well. Thus we can start with \( w = v \) and repeat the process, as needed, to get that (30.8.1) holds.

If \( u \in \Upsilon^+ \) and \( u \preceq v \), then

\[
(30.8.9) \quad \sigma(u) \in \Pi
\]

for every element \( \sigma \) of the Weyl group of \( \Phi \), because of (30.8.1) and (30.7.2). Conversely, if \( u_1 \in \Pi \), then there is an element \( \sigma_1 \) of the Weyl group of \( \Phi \) such that

\[
(30.8.10) \quad \sigma_1^{-1}(u_1) \in \Upsilon^+,
\]

as in Section 30.4. We also have that \( \sigma_1^{-1}(u_1) \in \Pi \), by (30.7.2), and that \( \sigma_1^{-1}(u_1) \preceq v \), because \( \Pi \) has highest weight \( v \), by hypothesis.

It follows that

\[
(30.8.11) \quad \Pi = \{ \sigma(u) : u \in \Upsilon^+, u \preceq v, \text{ and } \sigma \text{ is an element of the Weyl group of } \Phi \},
\]

as on p71 of [13]. In particular, this means that \( \Pi \) is uniquely determined by \( v \).
30.9 More on saturated sets

We continue with the same notations and hypotheses as in Sections 30.1 and 30.3 again, and to let $\preceq$ be defined on $V$ as in Section 30.5. We ask that $\Phi$ be reduced as a root system in $V$ in this section too.

Let $\Pi$ be a subset of $\Upsilon$ that is invariant under the Weyl group of $\Phi$, so that $\sigma(\Pi) = \Pi$ for every element $\sigma$ of the Weyl group. Suppose that for every $w \in \Pi \cap \Upsilon^+$ and $\alpha \in \Phi$ we have that

$$w - j \alpha \in \Pi$$

for every integer $j$ between 0 and $\lambda_\alpha(w)$. We would like to check that $\Pi$ is saturated as a subset of $\Upsilon$ under these conditions. This is related to a remark on p71 of [13].

To do this, let $v \in \Pi$ be given. Remember that there is an element $\sigma$ of the Weyl group such that $\sigma(v) \in \Upsilon^+$, as in Section 30.4. Note that $\sigma(v) \in \Pi$, by hypothesis. Let $\alpha \in \Phi$ be given, and remember that

$$\lambda_{\sigma(\alpha)}(\sigma(v)) = \lambda_\alpha(v),$$

as in (30.4.2).

If $j$ is an integer between 0 and $\lambda_{\sigma(\alpha)}(\sigma(v))$, then

$$\sigma(v) - j \sigma(\alpha) \in \Pi,$$

by hypothesis, and because $\sigma(\alpha) \in \Phi$. This implies that $v - j \alpha \in \Pi$, because $\sigma(\Pi) = \Pi$. Equivalently, this holds for every integer $j$ between 0 and $\lambda_\alpha(v)$, by (30.9.2), as desired.

Let $v \in \Upsilon^+$ be given, and let $\Pi = \Pi_\alpha$ be as in (30.8.11). We would like to verify that $\Pi$ is saturated as a subset of $\Upsilon$, as in Exercise 10 on p72 of [13]. Of course, $\Pi$ is invariant under the Weyl group of $\Phi$, by construction. Let $w \in \Pi \cap \Upsilon^+$ and $\alpha \in \Phi$ be given. It suffices to show that (30.9.1) holds for every integer $j$ between 0 and $\lambda_\alpha(w)$, as before.

By definition of $\Pi$, there is a $u \in \Upsilon^+$ and an element $\sigma$ of the Weyl group such that $w = \sigma(u)$ and $u \preceq v$. Under these conditions, we have that $u = w$, as in Section 30.4. Thus $w \preceq v$.

If $j \in \mathbb{Z}$, then $w - j \alpha \in \Upsilon$, and so there is an element $\sigma_j$ of the Weyl group such that

$$\sigma_j(w - j \alpha) \in \Upsilon^+,$$

as in Section 30.4. We would like to check that

$$\sigma_j(w - j \alpha) \preceq v$$

when $j$ is between 0 and $\lambda_\alpha(w)$.

Remember that $\sigma_j(w) \preceq w$, because $w \in \Upsilon^+$, as in Section 30.5. If

$$j \alpha \sigma_j(\alpha) \preceq 0,$$
then we get that
\[
\sigma_j(w - j \alpha) = \sigma_j(w) - j \sigma_j(\alpha) \preceq \sigma_j(w) \preceq w \preceq v.
\]

Observe that
\[
w - j \alpha = \sigma_\alpha(w) - (j - \lambda_\alpha(w)) \alpha,
\]
so that
\[
\sigma_j(w - j \alpha) = \sigma_j(\sigma_\alpha(w)) = (j - \lambda_\alpha(w)) \sigma_j(\alpha).
\]
As before, \(\sigma_j(\sigma_\alpha(w)) \preceq w\) because \(w \in Y^+\), as in Section 30.5. If
\[
(j - \lambda_\alpha(w)) \sigma_j(\alpha) \succeq 0,
\]
then it follows that
\[
\sigma_j(w - j \alpha) \preceq \sigma_j(\sigma_\alpha(w)) \preceq w \preceq v.
\]
Because \(\sigma_j(\alpha) \in \Phi\) we automatically have that \(\sigma_j(\alpha) \succeq 0\) or \(\sigma_j(\alpha) \preceq 0\). If
\(j\) is between 0 and \(\lambda_\alpha(w)\), then it is easy to see that one of \(j\) and \(j - \lambda_\alpha(w)\) is
greater than or equal to 0, and the other is less than or equal to 0. This implies
that (30.9.6) or (30.9.10) holds. It follows that (30.9.5) holds in either case, by
(30.9.7) and (30.9.11). This shows that \(\Pi\) is saturated in \(Y\), as desired.

It is easy to see that \(\Pi\) has highest weight \(v\), as on p71 of [13]. More precisely,
\(v \in \Pi\) by construction, because \(v \in Y^+\), by hypothesis. Let \(u \in Y^+\) with \(u \preceq v\)
be given, and let \(\sigma\) be an element of the Weyl group of \(\Phi\). Note that \(\sigma(u) \preceq u\),
as in Section 30.5. This implies that \(\sigma(u) \preceq v\), as desired.

### 30.10 Another look at saturated sets

Let us continue with the same notation and hypotheses as in the previous section. Also let \(v \in Y^+\) be given, and put
\[
\Pi^v = \{ u \in Y : \sigma(u) \preceq v \text{ for every element } \sigma \}
\]
of the Weyl group of \(\Phi\).

Observe that
\[
v \in \Pi^v,
\]
because \(v \in Y^+\), as in Section 30.5. Of course, \(\Pi^v\) is automatically invariant
under every element of the Weyl group of \(\Phi\).

Let \(\Pi_v = \Pi\) be the subset of \(Y\) defined as in (30.8.11) again, and let us check that
\[
\Pi^v = \Pi_v.
\]
By construction, every element of \(\Pi_v\) is of the form \(\tau(u)\), where \(u \in Y^+, u \preceq v,\)
and \(\tau\) is an element of the Weyl group of \(\Phi\). Let \(\sigma\) be any element of the Weyl
group of \(\Phi\), so that \(\sigma \circ \tau\) is an element of the Weyl group too. This implies that
\[
\sigma(\tau(u)) \preceq u,
\]
because \( u \in \mathbb{T}^+ \), as in Section 30.5. It follows that

\[(30.10.5) \quad \sigma(\tau(u)) \preceq v,\]

so that \( \tau(u) \in \Pi^v \).

Now let \( u \) be any element of \( \Pi^v \), and let \( \sigma \) be an element of the Weyl group of \( \Phi \) such that \( \sigma(u) \in \mathbb{T}^+ \), as in Section 30.4. We also have that \( \sigma(u) \preceq v \), by definition of \( \Pi^v \). This implies that \( u \in \Pi_v \), because the inverse of \( \sigma \) is in the Weyl group of \( \Phi \).

Let us verify directly that \( \Pi^v \) is saturated as a subset of \( \mathbb{T} \), as a variant of the argument in the previous section. Let \( u \in \Pi^v \) and \( \alpha \in \Phi \) be given, and let us check that

\[(30.10.6) \quad u - j \alpha \in \Pi^v\]

for every integer \( j \) between 0 and \( \lambda_\alpha(u) \). To do this, let an element \( \sigma \) of the Weyl group of \( \Phi \) be given, and let us show that

\[(30.10.7) \quad \sigma(u - j \alpha) \preceq v.\]

Of course, \( \sigma(u) \preceq v \), because \( u \in \Pi^v \). If

\[(30.10.8) \quad j \sigma(\alpha) \geq 0,\]

then

\[(30.10.9) \quad \sigma(u - j \alpha) = \sigma(u) - j \sigma(\alpha) \preceq \sigma(u) \preceq v.\]

We also have that

\[(30.10.10) \quad u - j \alpha = \sigma_\alpha(u) - (j - \lambda_\alpha(u)) \alpha,\]

so that

\[(30.10.11) \quad \sigma(u - j \alpha) = \sigma(\sigma_\alpha(u)) - (j - \lambda_\alpha(u)) \sigma(\alpha).\]

Note that \( \sigma(\sigma_\alpha(u)) \preceq v \), because \( u \in \Pi^v \) and \( \sigma \circ \sigma_\alpha \) is in the Weyl group of \( \Phi \). If

\[(30.10.12) \quad (j - \lambda_\alpha(u)) \sigma(\alpha) \geq 0,\]

then we get that

\[(30.10.13) \quad \sigma(u - j \alpha) \preceq \sigma(\sigma_\alpha(u)) \preceq v.\]

As before, we automatically have that \( \sigma(\alpha) \geq 0 \) or \( \sigma(\alpha) \leq 0 \), because \( \sigma(\alpha) \) is in \( \Phi \). We also have that one of \( j \) and \( j - \lambda_\alpha(u) \) is greater than or equal to 0, and the other is less than or equal to 0, because \( j \) is between 0 and \( \lambda_\alpha(u) \), by hypothesis. Using this, it is easy to see that (30.10.8) or (30.10.12) holds. In either case, we get that (30.10.7) holds, as desired. Note that \( \Pi^v \) has highest weight \( v \), by construction.
30.11 A property of saturated sets

Let us continue with the same notations and hypotheses as in Section 30.1 and 30.3, and let $\preceq$ be defined on $V$ as in Section 30.5. In this section, we continue to suppose that $\Phi$ is reduced as a root system in $V$. Let $(\cdot, \cdot)$ be an inner product on $V$ that is invariant under the Weyl group of $\Phi$.

Let $\Pi$ be a saturated subset of $\Upsilon$, with highest weight $v \in \Upsilon^+$. If $w \in \Pi$ and $\rho$ is as in Section 30.6, then Lemma C on p71 of [13] states that

$$(w + \rho, w + \rho) \leq (v + \rho, v + \rho),$$

with equality only when $w = v$. To see this, let $\sigma$ be an element of the Weyl group of $\Phi$ such that $u = \sigma(w) \in \Upsilon^+$, as in Section 30.4. Remember that

$$(w + \rho, w + \rho) \leq (u + \rho, u + \rho),$$

with equality only when $u = w$, as in Section 30.6. Thus it suffices to show that

$$(u + \rho, u + \rho) \leq (v + \rho, v + \rho),$$

with equality only when $u = v$.

Note that $u \in \Pi$, because $\sigma(\Pi) = \Pi$, as in Section 30.7. This means that $u \preceq v$, because $\Pi$ has highest weight $v$. Observe that

$$(v + \rho, v + \rho) - (u + \rho, u + \rho) = ((v + \rho, v + \rho) - (v + \rho, u + \rho)) + ((v + \rho, u + \rho) - (u + \rho, u + \rho)) = (v + \rho, v - u) + (v - u, u + \rho).$$

Remember that $\rho \in \Upsilon^+$, and in fact $\rho$ is strongly dominant, as in Section 30.6. This implies that $u + \rho$ and $v + \rho$ are strongly dominant too, because $u, v \in \Upsilon^+$.

It follows that each of the terms on the right side of (30.11.4) is greater than or equal to 0, with equality only when $u = v$, because $u \preceq v$. Combining this with (30.11.4), we get that (30.11.3) holds, with equality only when $u = v$, as desired.
Part V

Some representation theory
Chapter 31

Representations and semisimplicity

31.1 Diagonalizability and roots

Let \( k \) be a field of characteristic 0, and let \((A, \cdot , \cdot )\) be a Lie algebra over \( k \) of positive finite dimension. We would like to consider conditions on \( A \) like those in Section 29.4. More precisely, we shall start here with conditions related to diagonalizability, and consider additional conditions related to roots. In particular, these conditions hold when \( k \) is algebraically closed, and \( A \) is semisimple.

Suppose that \( A_0 \) is a Lie subalgebra of \( A \), and that every element of \( A_0 \) is \( \text{ad} \)-diagonalizable as an element of \( A \). This implies that \( A_0 \) is commutative as a Lie algebra, as in Section 17.1. Thus \( A_0 \) is contained in its centralizer \( C_A(A_0) \) in \( A \), and we ask that

\[
C_A(A_0) = A_0. \tag{31.1.1}
\]

Remember that this holds when \( k \) is algebraically closed, \( A \) is semisimple, and \( A_0 \) is a maximal toral subalgebra of \( A \), as in Section 17.4. Note that (31.1.1) implies that \( A_0 \neq \{0\} \), because \( A \neq \{0\} \).

Let \( A'_0 \) be the dual space of linear functionals on \( A_0 \), as a vector space over \( k \), as usual. If \( \alpha \in A'_0 \), then let \( A_\alpha \) be the set of \( x \in A \) such that \( \text{ad}_w(x) = \alpha(w)x \) for every \( w \in A_0 \), which is a linear subspace of \( A \). The condition (31.1.1) says exactly that \( A_\alpha \) is equal to \( A_0 \) when \( \alpha = 0 \).

Let \( A'_0 \) be the dual space of linear functionals on \( A_0 \), as a vector space over \( k \), as usual. If \( \alpha \in A'_0 \), then let \( A_\alpha \) be the set of \( x \in A \) such that \( \text{ad}_w(x) = \alpha(w)x \) for every \( w \in A_0 \), which is a linear subspace of \( A \). The condition (31.1.1) says exactly that \( A_\alpha \) is equal to \( A_0 \) when \( \alpha = 0 \).

Let \( \Phi = \Phi_{A_0} \) be the set of \( \alpha \in A_0' \) such that \( \alpha \neq 0 \) and \( A_\alpha \neq \{0\} \). Thus \( A \) corresponds to the direct sum of \( A_\alpha \), \( \alpha \in \Phi \cup \{0\} \), as a vector space over \( k \), as in Section 17.2. Remember that \( [A_\alpha, A_\beta] \subseteq A_{\alpha + \beta} \) for every \( \alpha, \beta \in A_0' \), as before.

Under the conditions mentioned so far, the center \( Z(A) \) of \( A \) as a Lie algebra is contained in \( A_0 \), and in fact is the intersection of the kernels of the elements of \( \Phi \). We ask that \( Z(A) = \{0\} \), which means that \( A'_0 \) is spanned by \( \Phi \), as a vector space over \( k \). This implies that \( \Phi \neq \emptyset \), because \( A_0 \neq \{0\} \).

We ask that for each \( \alpha \in \Phi \), \( A_\alpha \) have dimension one, as a vector space over \( k \). We also ask that for every \( \alpha \in \Phi \) there be an element \( h_\alpha \) of \( [A_\alpha, A_{-\alpha}] \subseteq A_0 \).
31.2. USING A LARGER FIELD

such that \( \alpha(h_{\alpha}) = 2 = 1 + 1 \) in \( k \). This determines \( h_{\alpha} \) uniquely, because \( k \)
has characteristic 0, and \( [A_{\alpha}, A_{-\alpha}] \) has dimension at most one as a vector space
over \( k \). Note that \( A \) is semisimple under these conditions, as in Section 17.15.
If \( \alpha, \beta \in \Phi \) and \( \alpha + \beta \in \Phi \), then we ask that \( [A_{\alpha}, A_{\beta}] = A_{\alpha+\beta} \).

Let \( E_Q \) be the linear subspace of \( A'_0 \), as a vector space over \( Q \), spanned by \( \Phi \).
This leads to a vector space \( E_R \) over \( R \), as in Section 17.13. More precisely,
\( E_Q \) corresponds to a linear subspace of \( E_R \), as a vector space over \( Q \), and any
basis for \( E_Q \), as a vector space over \( Q \), is a basis for \( E_R \), as a vector space over \( R \).

We ask that \( \Phi \) be a reduced root system in \( E_R \), and we let \( \Delta \) be a base for \( \Phi \).
In particular, this means that \( \Delta \) is a basis for \( E_R \), as a vector space over \( R \).
This implies that \( \Delta \) is a basis for \( E_Q \), as a vector space over \( Q \), because \( E_Q \) is
spanned by \( \Delta \), as a vector space over \( Q \).

Note that \( A'_0 \) is spanned by \( \Delta \), as a vector space over \( k \), because \( A'_0 \) is
spanned by \( \Phi \), as before. We ask that \( h_{\alpha}, \alpha \in \Delta \), form a basis for \( A'_0 \), as a vector space over \( k \).

Let \( \alpha \in \Phi \) be given, and let \( \sigma_{\alpha} \) be the symmetry on \( E_R \) with vector \( \alpha \)
that maps \( \Phi \) onto itself. Thus there is a linear functional \( \lambda_{\alpha} \) on \( E_R \) so that \( \sigma_{\alpha} \) minus
the identity mapping on \( E_R \) is equal to \(-\lambda_{\alpha} \) times \( \alpha \). We ask that

\[
(31.1.2) \quad \beta(h_{\alpha}) = \lambda_{\alpha}(\beta)
\]

for every \( \beta \in \Phi \), where the right side is an integer by the definition of a root
system, which may be considered as an element of \( k \).

If \( \Psi_0 \subseteq \Phi \cup \{0\} \), then \( A(\Psi_0) \) denotes the linear subspace of \( A \) spanned by
\( A_{\alpha}, \alpha \in \Psi_0 \), as in Section 22.4. Let \( \Phi^+ = \Phi^{\Delta^+} \) be the set of positive roots in \( \Phi \)
with respect to \( \Delta \), which is to say the elements of \( \Phi \) that can be expressed as
linear combinations of elements of \( \Delta \) with nonnegative coefficients. Put

\[
(31.1.3) \quad B_\Delta = A(\Phi^+ \cup \{0\}),
\]

which is the standard Borel subalgebra of \( A \) associated to \( A_0 \) and \( \Delta \), as in
Section 22.12.

31.2 Using a larger field

Let \( k, k_1 \) be fields, with \( k \subseteq k_1 \). If \( V \) is a vector space over \( k \), then we can get
a vector space \( \tilde{V} \) over \( k_1 \) from \( V \) in a well-known way. Of course, \( \tilde{V} \) may also
be considered as a vector space over \( k \), and \( V \) may be considered as a linear
subspace of \( \tilde{V} \), as a vector space over \( k \). If \( \{v_j\}_{j \in I} \) is a basis for \( V \) as a vector
space over \( k \), then \( \{v_j\}_{j \in I} \) corresponds to a basis for \( \tilde{V} \) as a vector space over \( k_1 \) too.
In particular, the dimension of \( \tilde{V} \), as a vector space over \( k_1 \), is the same
as the dimension of \( V \), as a vector space over \( k \).

One can use a basis for \( V \) to obtain \( \tilde{V} \) in this way, and verify that \( \tilde{V} \) does not
depend on the choice of basis for \( V \), up to isomorphic equivalence. In particular,
every element of $V$ can be expressed as a finite linear combination of elements of $V$, with coefficients in $k_1$. Any collection of linearly independent vectors in $V$, as a vector space over $k$, is linearly independent in $V$ as well, as a vector space over $k_1$.

If $W_1$ is a vector space over $k_1$, then $W_1$ may be considered as a vector space over $k$. Every linear mapping from $V$ into $W_1$ as a vector space over $k$ has a unique extension to a linear mapping from $V$ into $W_1$, as vector spaces over $k_1$. Of course, the restriction of any linear mapping from $V$ into $W_1$, as a vector space over $k_1$, to $V$ is linear as a mapping into $W_1$, as a vector space over $k$.

If $Z$ is a linear subspace of $V$, then one can get a vector space $\tilde{Z}$ over $k_1$ in the same way. More precisely, $\tilde{Z}$ corresponds to a linear subspace of $V$, whose intersection with $V$ is $Z$. If $V$ is the direct sum of some family of vector spaces over $k$, then $\tilde{V}$ is the direct sum of the corresponding family of vector spaces over $k_1$.

Let $Z$ be any vector space over $k$, and let $\tilde{Z}$ be the corresponding vector space over $k_1$. If $W_1$ is a vector space over $k_1$ again, then any mapping from $V \times Z$ into $W_1$ that is bilinear over $k$ has a unique extension to a mapping from $V \times \tilde{Z}$ into $W_1$, as a vector space over $k_1$.

Let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$, so that $A$ is a vector space over $k$ in particular. Thus we can get a corresponding vector space $\tilde{A}$ over $k_1$ as before. The Lie bracket $[\cdot, \cdot]_A$ has a unique extension to a mapping $[\cdot, \cdot]_{\tilde{A}}$ from $\tilde{A} \times \tilde{A}$ into $\tilde{A}$ that is bilinear over $k_1$, as in the previous paragraph. One can check that $\tilde{A}$ is a Lie algebra over $k_1$ with respect to $[\cdot, \cdot]_{\tilde{A}}$. Note that $\tilde{A}$ may be considered as a Lie algebra over $k$, and that $A$ is a Lie subalgebra of $\tilde{A}$, as a Lie algebra over $k$.

If $B$ is a Lie subalgebra of $A$, then $B$ is a linear subspace of $A$ in particular, and $\tilde{B}$ corresponds to a linear subspace of $\tilde{A}$, as a vector space over $k_1$. It is easy to see that $\tilde{B}$ is a Lie subalgebra of $\tilde{A}$, as a Lie algebra over $k_1$.

31.3 Related spaces of linear mappings

Let $k$, $k_1$ be fields with $k \subseteq k_1$ again, let $V$ be a vector space over $k$, and let $W_1$ be a vector space over $k_1$. Thus $W_1$ may also be considered as a vector space over $k$, and we let $\mathcal{L}_k(V, W_1)$ be the space of linear mappings from $V$ into $W_1$, as a vector space over $k$. Of course, the space of all mappings from any nonempty set into $W_1$ may be considered as a vector space over $k_1$, with respect to pointwise addition and scalar multiplication. It is easy to see that $\mathcal{L}_k(V, W_1)$ is a linear subspace of the space of all mappings from $V$ into $W_1$, as a vector space over $k_1$. This means that $\mathcal{L}_k(V, W_1)$ may be considered as a vector space
31.3. RELATED SPACES OF LINEAR MAPPINGS

over $k_1$.

There is a natural mapping from $L_k(V,W_1)$ into the space $L_{k_1}((V,W_1)$) of all linear mappings from $V$ into $W_1$, as vector spaces over $k_1$. This mapping sends an element of $L_k(V,W_1)$ to its unique extension to an element of $L_{k_1}((V,W_1)$, as before. Similarly, there is a natural mapping from $L_{k_1}((V,W_1)$) into $L_k(V,W_1)$, which sends an element of $L_{k_1}((V,W_1)$ to its restriction to $V$. Of course, these two mappings are inverses of each other. These mappings are also linear over $k_1$, so that we get

\begin{equation}
\text{(31.3.1)}
\end{equation}

\text{a natural isomorphism from } L_k(V,W_1) \text{ onto } L_{k_1}((V,W_1),
\text{as vector spaces over } k_1.

Let $W$ be another vector space over $k$, and let $\tilde{W}$ be the corresponding vector space over $k_1$ that contains $W$ as a linear subspace. A linear mapping from $V$ into $W$, as vector spaces over $k$, may be considered as a linear mapping from $V$ into $\tilde{W}$, as vector spaces over $k$. This may be extended to a unique linear mapping from $V$ into $\tilde{W}$, as vector spaces over $k_1$. We can take $W_1 = \tilde{W}$ in (31.3.1), to get

\begin{equation}
\text{(31.3.2)}
\end{equation}

\text{a natural isomorphism from } L_k(V,\tilde{W}) \text{ onto } L_{k_1}((V,\tilde{W}),
\text{as vector spaces over } k_1.

We may also consider $\tilde{W}$ and $L_k(V,\tilde{W})$ as vector spaces over $k$, and the space $L_k(V,W)$ of all linear mappings from $V$ into $W$, as vector spaces over $k$, is a linear subspace of $L_k(V,\tilde{W})$, as a vector space over $k$. Thus

\begin{equation}
\text{(31.3.3)}
\end{equation}

$L_k(V,W)$ corresponds to a linear subspace of $L_{k_1}((V,\tilde{W}),$
\text{as a vector space over } k,

using the natural isomorphism mentioned in (31.3.2).

Suppose that $\{v_i\}_{i \in I}$ is a basis for $V$ as a vector space over $k$, which may also be considered as a basis for $V$ as a vector space over $k_1$. We may as well suppose that $I \neq \emptyset$, which is to say that $V \neq \{0\}$. Let $c(I,W)$, $c(I,\tilde{W})$ be the spaces of functions on $I$ with values in $W$, $\tilde{W}$, respectively. These are vector spaces over $k$, $k_1$, respectively, with respect to pointwise addition and scalar multiplication.

A linear mapping from $V$ into $W$ is determined by its values on the basis vectors, and any element of $c(I,W)$ determines such a linear mapping in this way. This defines

\begin{equation}
\text{(31.3.4)}
\end{equation}

\text{an isomorphism from } L_k(V,W) \text{ onto } c(I,W),
\text{as vector spaces over } k.

Similarly, we get

\begin{equation}
\text{(31.3.5)}
\end{equation}

\text{an isomorphism from } L_k(V,\tilde{W}) \text{ onto } c(I,\tilde{W}),
\text{as vector spaces over } k_1.
In the same way, we obtain

\[(31.3.6) \quad \text{an isomorphism from } \mathcal{L}_{k_1}(\tilde{V}, \tilde{W}) \text{ onto } c(I, \tilde{W}), \]

as vector spaces over \(k_1\).

Of course, (31.3.5) and (31.3.6) correspond to each other as in (31.3.2). Note that \(c(I, W)\) may be considered as a vector space over \(k\) by considering \(\tilde{W}\) as a vector space over \(k_1\).

\[(31.3.7) \quad c(I, W) \text{ is a linear subspace of } c(I, \tilde{W}), \quad \text{as a vector space over } k, \]

because \(W\) is a linear subspace of \(\tilde{W}\), as a vector space over \(k\). This corresponds to the fact that \(\mathcal{L}_k(V, W)\) may be considered as a linear subspace of \(\mathcal{L}_{k_1}(V, \tilde{W})\), as a vector space over \(k\), using (31.3.4) and (31.3.5). Similarly, (31.3.3) corresponds to (31.3.7), using (31.3.4) and (31.3.6).

Let \(V', \tilde{V}'\) be the duals of \(V, \tilde{V}\), as vector spaces over \(k, k_1\), respectively. Thus \(V'\) consists of the linear functionals on \(V\), which are linear mappings from \(V\) into \(k\), and \(\tilde{V}'\) consists of mappings from \(\tilde{V}\) into \(k_1\) that are linear over \(k_1\). Of course, \(W = k\) may be considered as a one-dimensional vector space over \(k\), for which we get \(\tilde{W} = k_1\), as a one-dimensional vector space over \(k_1\).

The space \(\mathcal{L}_k(V, k_1)\) of linear mappings from \(V\) into \(k_1\), as a vector space over \(k\), may be considered as a vector space over \(k_1\), as in Section 31.2. As before, \(\mathcal{L}_k(V, k_1)\) contains \(V'\) as a linear subspace. This means that

\[(31.3.8) \quad V' \text{ corresponds to a linear subspace of } \tilde{V}', \quad \text{as a vector space over } k, \]

in a natural way, as in (31.3.3).

Suppose that \(V\) has positive finite dimension \(n\), and let \(v_1, \ldots, v_n\) be a basis for \(V\). This may also be considered as a basis for \(\tilde{V}\), as a vector space over \(k_1\), as before. Using this basis, we can identify \(V'\) with \(k^n\), and \(\tilde{V}'\) with \(k_1^n\). One can use this to see that the vector space \((\tilde{V}')\) over \(k_1\) associated to \(V'\) can be identified with \(\tilde{V}'\).

### 31.4 Larger fields and toral subalgebras

Let \(k, k_1\) be fields with \(k \subseteq k_1\) again, and let \((A, [\cdot, \cdot]_A)\) be a Lie algebra over \(k\) of positive finite dimension. This leads to a vector space \(\tilde{A}\) over \(k_1\) of the same dimension, with \(A\) a linear subspace of \(\tilde{A}\) as a vector space over \(k\), as in Section 31.2. As before, \([\cdot, \cdot]_A\) has a unique extension to a mapping \([\cdot, \cdot]_{\tilde{A}}\) from \(\tilde{A} \times \tilde{A}\) into \(\tilde{A}\) that is bilinear over \(k_1\), and which makes \((\tilde{A}, [\cdot, \cdot]_{\tilde{A}})\) into a Lie algebra over \(k_1\).

Suppose that \(A_0\) is a Lie subalgebra of \(A\) such that every element of \(A_0\) is ad-diagonalizable as an element of \(A\), so that \(A_0\) is commutative as a Lie
31.4. LARGER FIELDS AND TORAL SUBALGEBRAS

algebra, as in Section 17.1. If $\alpha$ is an element of the dual $A'_0$ of $A_0$, as a vector space over $k$, then let $A_\alpha$ be the set of $x \in A$ such that $\text{ad}_{A_\alpha}(x) = \alpha(w)x$ for every $w \in A_0$, as before. Suppose that $A_0$ is its own centralizer in $A$, as in (31.1.1), so that $A_0$ is the same as $A_\alpha$ with $\alpha = 0$.

Let $\Phi = \Phi_{A_0}$ be the set of $\alpha \in A'_0$ such that $\alpha \neq 0$ and $A_\alpha \neq \{0\}$, as before. Remember that $A$ corresponds to the direct sum of the linear subspaces $A_\alpha$, $\alpha \in \Phi \cup \{0\}$, as a vector space over $k$, as in Section 17.2. If $\alpha \in A'_0$, so that $A_\alpha$ is a linear subspace of $A$, then let $(A_\alpha)$ be the linear subspace of $\widetilde{A}$ as a vector space over $k_1$ obtained from $A_\alpha$ as before. Observe that $\widetilde{A}$ corresponds to the direct sum of $(A_\alpha)$, $\alpha \in \Phi \cup \{0\}$, as a vector space over $k_1$, as in Section 31.2.

More precisely, $(\widetilde{A}_0)$ is a Lie subalgebra of $\widetilde{A}$, as a Lie algebra over $k_1$. If $\alpha \in A'_0$, then let $\hat{\alpha}$ be the unique extension of $\alpha$ to a mapping from $(\widetilde{A}_0)$ into $k_1$ that is linear over $k_1$, as in Section 31.2.

In particular, the elements of $(\widetilde{A}_0)$ are ad-diagonalizable in $\widetilde{A}$. If $\beta \in (\widetilde{A}_0)'$, then we can take $\tilde{A}_\beta$ to be the set of $x \in \tilde{A}$ such that $\text{ad}_{\tilde{A}_\alpha}(x) = \beta(w)x$ for every $w \in \widetilde{A}_0$, as usual. If $\alpha \in A'_0$ and $\hat{\alpha}$ is the corresponding element of $(\widetilde{A}_0)'$, then

\begin{equation}
\text{ad}_{\tilde{A}_\alpha}(x) = [w, x]_{\tilde{A}} = \hat{\alpha}(w)x
\end{equation}

for every $w \in (\widetilde{A}_0)$ and $x \in (\tilde{A}_\alpha)$, because this holds when $w \in A_0$ and $x \in A_\alpha$, by construction.

In particular, the elements of $(\widetilde{A}_0)$ are ad-diagonalizable in $\widetilde{A}$. If $\beta \in (\widetilde{A}_0)'$, then we can take $\tilde{A}_\beta$ to be the set of $x \in \tilde{A}$ such that $\text{ad}_{\tilde{A}_\alpha}(x) = \beta(w)x$ for every $w \in \widetilde{A}_0$, as usual. Thus

\begin{equation}
\{\hat{\alpha} : \alpha \in \Phi\} \subseteq \Phi,
\end{equation}

by (31.4.2). Let $\tilde{\Phi} = \tilde{\Phi}_{(\widetilde{A}_0)}$ be the set of $\beta \in (\widetilde{A}_0)'$ such that $\beta \neq 0$ and $\tilde{A}_\beta \neq \{0\}$, as usual. Thus

\begin{equation}
\{\hat{\alpha} : \alpha \in \Phi\} \subseteq \Phi,
\end{equation}

by (31.4.2).

In fact,

\begin{equation}
\tilde{\Phi} = \{\hat{\alpha} : \alpha \in \Phi\},
\end{equation}

and

\begin{equation}
\tilde{A}_\alpha = (A_\alpha)
\end{equation}

for every $\alpha \in A'_0$. This follows from (31.4.1) and the fact that $\tilde{A}$ corresponds to the direct sum of $(\widetilde{A}_\alpha)$, $\alpha \in \Phi \cup \{0\}$, as a vector space over $k_1$. In particular, if $\alpha = 0$, then $\hat{\alpha} = 0$ as an element of $(\widetilde{A}_0)'$, and we get that

\begin{equation}
\tilde{A}_0 = (\widetilde{A}_0).
\end{equation}

Note that $(\widetilde{A}_0)$ is commutative as a Lie algebra over $k_1$, and that it is its own centralizer in $\tilde{A}$, by (31.4.6).
\[ (31.4.7) \quad [(A_\alpha), (A_{\beta})] = (A_{\alpha+\beta}) \]

in \( \tilde{A} \).

Every element of \( (\tilde{A}_0) \) can be expressed as a linear combination of elements of the form \( \tilde{\alpha}, \alpha \in \tilde{A}_0 \). Suppose that \( A_0' \) is spanned by \( \Phi \), as a vector space over \( k \). This means that the center \( Z(A) \) of \( A \) as a Lie algebra over \( k \) is trivial, as in Section 31.1. Under these conditions, we get that \( (\tilde{A}_0) \) is spanned by \( \tilde{\Phi} \), as a vector space over \( k_1 \). This means that \( Z(\tilde{A}) = \{0\} \), as before.

Suppose that for each \( \alpha \in \Phi \), \( A_\alpha \) has dimension one, as a vector space over \( k \). This implies that \( (\tilde{A}_0) \) has dimension one, as a vector space over \( k_1 \). Suppose also that for every \( \alpha \in \Phi \) there is an element \( h_\alpha \) of \( [A_\alpha, A_{-\alpha}] \subseteq A_0 \) such that \( \alpha(h_\alpha) = 2 \). Of course, \( h_\alpha \) may also be considered as an element of \( [(A_{\tilde{\alpha}}), (A_{-\tilde{\alpha}})] \) in \( \tilde{A} \).

Suppose that \( k \) has characteristic 0, so that \( k_1 \) has characteristic 0 as well. The linear subspace \( E_\Phi \) of \( A_0' \), as a vector space over \( \mathbb{Q} \), corresponds exactly to the linear subspace of \( (\tilde{A}_0) \) spanned by \( \tilde{\Phi} \), as a vector space over \( \mathbb{Q} \), because of (31.4.4). This means that the associated vector space \( E_{\mathbb{R}} \) over \( \mathbb{R} \) corresponds exactly to the analogous vector space over \( \mathbb{R} \) for \( (\tilde{A}_0) \).

Suppose that \( \Phi \) is a reduced root system in \( E_{\mathbb{R}} \), so that \( \tilde{\Phi} \) is a reduced root system in the analogous vector space over \( \mathbb{R} \) for \( (\tilde{A}_0) \). Let \( \Delta \) be a base for \( \Phi \), which means that \( \tilde{\Delta} = \{\tilde{\alpha} : \alpha \in \Delta\} \) is a base for \( \tilde{\Phi} \). If \( h_\alpha, \alpha \in \Delta \), is a basis for \( A_0 \), as a vector space over \( k \), then this is also a basis for \( (\tilde{A}_0) \), as a vector space over \( k_1 \).

If \( Z(A) = \{0\} \), then \( A_0' \) is spanned by \( \Phi \) as a vector space over \( k \), and \( (\tilde{A}_0) \) is spanned by \( \tilde{\Phi} \) as a vector space over \( k_1 \), as before. In this case, we get that \( A_0' \) is spanned by \( \Delta \) as a vector space over \( k \), and that \( (\tilde{A}_0) \) is spanned by \( \tilde{\Delta} \) as a vector space over \( k_1 \). If \( \Delta \) is a basis for \( A_0' \) as a vector space over \( k \), then \( \tilde{\Delta} \) is a basis for \( (\tilde{A}_0) \), as a vector space over \( k_1 \).

If \( \alpha, \beta \in \Phi \), then one may ask that \( \beta(h_\alpha) \) can be expressed in terms of \( \Phi \) as a root system as in (31.1.2). This implies the analogous condition for \( \tilde{\Phi} \).

If \( \Psi_0 \subseteq \Phi \cup \{0\} \), then \( A(\Psi_0) \) is the linear subspace of \( A \) spanned by \( A_{\alpha}, \alpha \in \Psi_0 \), as in Section 31.1. Similarly, if \( \Psi_1 \subseteq \tilde{\Phi} \cup \{0\} \), then let \( \tilde{A}(\Psi_1) \) be the linear subspace of \( \tilde{A} \), as a vector space over \( k_1 \), spanned by \( \tilde{A}_\gamma, \gamma \in \Psi_1 \), as before. If \( \Psi_0 \subseteq \Phi \cup \{0\} \), then
\[ (31.4.8) \quad \Psi_1 = \{\tilde{\alpha} : \alpha \in \Psi_0\} \]
is contained in \( \tilde{\Phi} \cup \{0\} \), and
\[ (31.4.9) \quad \tilde{A}(\Psi_1) = (A(\Psi_0)), \]
31.5. WEIGHTS AND MODULES

Let \( \Phi^+ = \Phi^\Delta + \) be the set of positive roots in \( \Phi \) with respect to \( \Delta \), and let \( \tilde{\Phi}^+ = \tilde{\Phi}_{\tilde{\Delta}} + \) be the set of positive roots in \( \tilde{\Phi} \) with respect to \( \tilde{\Delta} \). Observe that

\[
(31.4.10) \quad \tilde{\Phi}^+ = \{ \tilde{\beta} : \beta \in \Phi^+ \}.
\]

Remember that, under the conditions on \( A \) and \( A_0 \) considered in Section 31.1, \( B_\Delta = A(\Phi^+ \cup \{0\}) \) is the standard Borel subalgebra of \( A \) associated to \( A_0 \) and \( \Delta \), as in Section 22.12. In this case, \( \tilde{A} \) and \( (\tilde{A}_0) \) satisfy the analogous conditions over \( k_1 \), and

\[
(31.4.11) \quad \tilde{B}_\tilde{\Delta} = \tilde{A}(\tilde{\Phi}^+ \cup \{0\})
\]

is the standard Borel subalgebra of \( \tilde{A} \), as a Lie algebra over \( k_1 \), associated to \( (\tilde{A}_0) \) and \( \tilde{\Delta} \). We also have that

\[
(31.4.12) \quad \tilde{B}_\Delta = (\tilde{B}_\Delta),
\]

by (31.4.9) and (31.4.10).

31.5 Weights and modules

Let us return to the same notation and hypotheses as in Section 31.1. Let \( V \) be a vector space over \( k \) which is a module over \( A \), as a Lie algebra over \( k \). If \( \mu \in A_0' \), then put

\[
(31.5.1) \quad V_\mu = \{ v \in V : w \cdot v = \mu(w) v \text{ for every } w \in A_0 \},
\]

which is a linear subspace of \( V \). An element of \( V_\mu \) is said to have weight \( \mu \) with respect to \( A_0 \), as on p56 of [23].

The elements of \( V_\mu \) may also be called eigenvectors of \( A_0 \) in \( V \) with weight \( \mu \), as on p57 of [24]. More precisely, Chapter VII of [24] deals with representations of \( sl_n(k), n \geq 2 \). However, all but one of the results discussed there extend to arbitrary semisimple Lie algebras, as in Remark 1 on p61 of [24].

The dimension of \( V_\mu \) is called the multiplicity of \( \mu \) in \( V \), as on p56 of [23], and Definition 2.3 on p58 of [24]. In particular, \( \mu \) is said to be a weight of \( V \) with respect to \( A_0 \) when \( V_\mu \neq \{0\} \), as on p107 of [13], p56 of [23], and Definition 2.3 on p58 of [24]. In this case, \( V_\mu \) may be called a weight space of \( V \) with respect to \( \mu \), as on p107 of [13].

Let \( \alpha, \mu \in A_0' \) be given, and suppose that \( x \in A_\alpha, v \in V_\mu \). If \( w \in A_0 \), then

\[
(31.5.2) \quad w \cdot (x \cdot v) = x \cdot (w \cdot v) + ([w, x]_A) \cdot v = \mu(w) (x \cdot v) + \alpha(w) (x \cdot v).
\]

This shows that

\[
(31.5.3) \quad x \cdot v \in V_{\mu + \alpha},
\]

as in part (a) of the lemma on p107 of [13], part (a) of Proposition 1 on p56 of [23], and Proposition 2.1 on p57 of [24]. Of course, this can be seen more directly when \( \alpha = 0 \), and it is trivial when \( \alpha \not\in \Phi \cup \{0\} \), in which case \( A_\alpha = \{0\} \).
656  

CHAPTER 31. REPRESENTATIONS AND SEMISIMPLICITY

Let \( \hat{V} \) be the linear subspace of \( V \) spanned by \( V_\mu, \mu \in \Lambda'_0 \). In fact,

\[
\hat{V} \text{ corresponds to the direct sum of } V_\mu, \mu \in \Lambda'_0,
\]
as a vector space over \( k \).

This means that if \( \mu_1, \ldots, \mu_r \) are finitely many distinct elements of \( \Lambda'_0 \), \( v_j \in V_{\mu_j} \) for each \( j = 1, \ldots, r \), and

\[
\sum_{j=1}^{r} v_j = 0,
\]
then \( v_j = 0 \) for every \( j = 1, \ldots, r \). This can be obtained from the fact that eigenvectors of a linear mapping with distinct eigenvalues are linearly independent. This corresponds to the first part of part (b) of the lemma on p107 of [13], the first part of part (b) of Proposition 1 on p56 of [23], and part of Proposition 2.2 on p57 of [24].

It is easy to see that \( \hat{V} \) is a submodule of \( V \), as a module over \( A \), using (31.5.3). This corresponds to the second part of part (b) of the lemma on p107 of [13], the second part of part (b) of Proposition 1 on p56 of [23], and a remark in the proof of Proposition 2.2 on p57 of [24].

As an example, consider \( A \) as a module over itself, with respect to the adjoint representation. In this case, the weights are the roots \( \alpha \in \Phi \), with the corresponding weight spaces \( A_{\alpha} \) of dimension 1, and 0. This is Example (1) on p107 of [13], and Example 2.4 of p58 of [24].

## 31.6 Primitive or maximal vectors

Let us continue to use the same notation and hypotheses as in Section 31.1. In particular, \( \Delta \) is a base for \( \Phi \) as a root system, and \( \Phi^+ \) is the set of positive roots in \( \Phi \) with respect to \( \Delta \). Let \( V \) be a vector space over \( k \) that is a module over \( A \) as a Lie algebra over \( k \) again.

If \( v \in V \) and \( E \subseteq A \), then put

\[
E \cdot v = \{ x \cdot v : x \in E \},
\]
which is a linear subspace of \( V \) when \( E \) is a linear subspace of \( A \). Suppose that

\[
E_1 \cdot v = E_2 \cdot v = \{ 0 \}
\]
for some linear subspaces \( E_1, E_2 \) of \( A \). If \( x_1 \in E_1 \) and \( x_2 \in E_2 \), then we get

\[
([x_1, x_2]_A) \cdot v = x_1 \cdot (x_2 \cdot v) - x_2 \cdot (x_1 \cdot v) = 0.
\]
This means that

\[
([E_1, E_2]) \cdot v = \{ 0 \}
\]
when (31.6.2) holds. If \( \alpha, \beta \in \Phi \) satisfy \( A_\alpha \cdot v = A_\beta \cdot v = \{ 0 \} \) and \( \alpha + \beta \in \Phi \), then it follows that

\[
(A_{\alpha + \beta}) \cdot v = \{ 0 \},
\]
31.6. PRIMITIVE OR MAXIMAL VECTORS

because \([A_\alpha, A_\beta] = A_{\alpha + \beta}\).

If
\[
A_\alpha \cdot v = \{0\}
\]
for every \(\alpha \in \Delta\),
then
\[
A_\gamma \cdot v = \{0\}
\]
for every \(\gamma \in \Phi^+\).

This is implicit in a remark on p108 of [13], and on p57 of [23]. To see this, let \(\gamma \in \Phi^+\) be given, and remember that \(\gamma\) can be expressed as \(\sum_{j=1}^r \alpha_j\), where \(\alpha_j \in \Delta\) for each \(j = 1, \ldots, r\), and \(\sum_{j=1}^l \alpha_j \in \Phi\) for every \(l = 1, \ldots, r\), as in Section 19.12. Under these conditions, one can check that
\[
(A \left(\sum_{j=1}^l \alpha_j\right)) \cdot v = \{0\}
\]
for every \(l = 1, \ldots, r\), using (31.6.5).

Let \(\mu\) be a linear functional on \(A_0\). We say that \(v \in V_\mu\) is a primitive or maximal vector of weight \(\mu\) in \(V\) if \(v \neq 0\) and (31.6.6) or equivalently (31.6.7) holds, as on p108 of [13], and p57 of [23]. We may simply say that \(v \in V\) is a primitive or maximal vector in \(V\) if \(v\) is primitive or maximal of weight \(\mu\) for some \(\mu \in A_0^*\).

Remember that \(B_\Delta\) is the standard Borel subalgebra of \(A\) associated to \(A_0\) and \(\Delta\), as in Section 31.1. Remember also that \(N_\Delta = A(\Phi^+)\) is a Lie subalgebra of \(A\) that is nilpotent as a Lie algebra over \(k\) and whose elements are ad-nilpotent in \(A\), and that \([B_\Delta, B_\Delta] = N_\Delta\), as in Sections 22.6 and 24.10. By construction, \(B_\Delta\) corresponds to the direct sum of \(A_0\) and \(N_\Delta\), as a vector space over \(k\).

Let us say that \(v \in V\) is an eigenvector of \(x \in A\) if \(x \cdot v\) is the multiple of \(v\) by an element of \(k\), which is the corresponding eigenvalue. Of course, this is the same as saying that \(v\) is an eigenvector for the linear mapping corresponding to \(x\) on \(V\), as a module over \(A\). If \(E \subseteq A\) and \(v\) is an eigenvector of every \(x \in E\), then \(v\) is said to be an eigenvector of \(E\). If \(v\) is an eigenvector of \(x_1, x_2 \in A\) with eigenvalues \(c_1, c_2 \in k\), then
\[
([x_1, x_2]_A) \cdot v = x_1 \cdot (x_2 \cdot v) - x_2 \cdot (x_1 \cdot v)
\]
\[
= c_2 (x_1 \cdot v) - c_1 (x_2 \cdot v) = c_1 c_2 v - c_1 c_2 v = 0.
\]

This means that \(v\) is an eigenvector of \([x_1, x_2]_A\), with eigenvalue 0.

If \(v \in V\) is a eigenvector of \(B_\Delta\), then \(v\) is an eigenvector of every element of \(N_\Delta\) with eigenvalue 0, because \([B_\Delta, B_\Delta] = N_\Delta\). This is the same as saying that (31.6.7) holds. More precisely, \(v \in V\) is an eigenvector of \(B_\Delta\) if and only if \(v\) is an eigenvector of \(A_0\) and (31.6.7) holds, as in Proposition 2.5 on p58 of [24].

Note that a nonzero \(v \in V\) is an eigenvector of \(A_0\) if and only if \(v \in V_\mu\) for some \(\mu \in A_0^*\), and that \(\mu\) is unique. Thus a nonzero \(v \in V\) is primitive or maximal if and only if it is an eigenvector of \(A_0\) and satisfies (31.6.7), or equivalently \(v\) is an eigenvector of \(B_\Delta\). This corresponds to Definition 2.6 on p58 of [24], and is also mentioned on p57 of [23].
31.7 Representations and $UA$

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. Also let $V$ be a module over $k$, which is a module over $A$, as a Lie algebra over $k$. Suppose that $UA$ is a universal enveloping algebra of $A$ with multiplicative identity element $e = e_{UA}$ and associated mapping $i = i_{UA}$ from $A$ into $UA$, as in Section 25.4.

Remember that the space $\text{Hom}_k(V, V)$ of homomorphisms from $V$ into itself, as a module over $k$, is an associative algebra over $k$ with respect to composition of mappings. In particular, $\text{Hom}_k(V, V)$ is a Lie algebra over $k$, with respect to the commutator bracket. To say that $V$ is a module over $A$ as a Lie algebra over $k$ means that we have a Lie algebra homomorphism from $A$ into $\text{Hom}_k(V, V)$.

Because $UA$ is a universal enveloping algebra of $A$, there is a corresponding homomorphism from $UA$ into $\text{Hom}_k(V, V)$, as associative algebras over $k$, that sends $e$ to the identity mapping on $V$. This defines a representation of $UA$, as an associative algebra over $k$, on $V$. Equivalently, $V$ may be considered as a (left) module over $UA$, as an associative algebra over $k$.

If $x \in A$ and $v \in V$, then the action of $x$ on $v$ may be denoted $x \cdot v$, as usual. Similarly, if $u \in UA$ and $v \in V$, then let the action of $u$ on $v$ be denoted $u \cdot v$. It should be clear in practice which is intended, and the two are compatible in the sense that $x \cdot v$ is the same as $i(x) \cdot v$ for every $x \in A$.

Let $v \in V$ be given, and define $E \cdot v \subseteq V$ as in (31.6.1) when $E \subseteq A$. This is a submodule of $V$, as a module over $k$, when $E$ is a submodule of $A$, as a module over $k$. If $E \subseteq UA$, then $E \cdot v$ can be defined as a subset of $V$ analogously, and is a submodule of $V$ as a module over $k$ when $E$ is a submodule of $UA$ as a module over $k$. Note that $v \in E \cdot v$ when $e \in E \subseteq UA$.

Clearly every submodule of $V$, as a module over $UA$, is a submodule of $V$, as a module over $A$. Conversely, any submodule of $V$ as a module over $A$ is a submodule of $V$ as a module over $UA$, because $UA$ is generated by $e$ and $i(A)$, as an associative algebra over $k$.

If $v \in V$, then it is easy to see that $(UA) \cdot v$ is a submodule of $V$, as a module over $UA$, that contains $v$, and in fact it is the smallest such submodule of $V$.

Equivalently, $(UA) \cdot v$ is a submodule of $V$, as a module over $A$, that contains $v$, and is the smallest such submodule.

Remember that $i$ is injective as a mapping from $A$ into $UA$ when $A$ is free as a module over $k$, as in Section 25.12. In this case, we may identify $A$ with $i(A) \subseteq UA$. In particular, this condition holds when $k$ is a field.

31.8 Spanning standard cyclic modules

Let us return now to the notation and hypotheses in Section 31.1, so that in particular $k$ is a field of characteristic 0, and $(A, [\cdot, \cdot]_A)$ is a Lie algebra over $k$ of positive finite dimension. Let $V$ be a module over $A$ again, as a Lie algebra over $k$, and let $UA$ be a universal enveloping algebra of $A$, so that $V$ may also be considered as a module over $UA$, as an associative algebra over $k$. Suppose
that \( v \in V \) is a primitive or maximal vector of weight \( \mu \in A'_0 \), as in Section 31.6, and that

\[
(31.8.1) \quad V = (UA) \cdot v.
\]

In this case, \( V \) is said to be \textit{standard cyclic} of weight \( \mu \), as on p108 of [13]. Of course, if (31.8.1) does not hold, \((UA) \cdot v\) is a submodule of \( V \), as a module over \( A \), which is standard cyclic of weight \( \mu \) as a module over \( A \).

If \( \alpha \in \Phi^+ \), then let \( x_\alpha \) be a nonzero element of \( A_\alpha \), so that \( A_\alpha \) consists of multiples of \( x_\alpha \) by elements of \( k \). Also let \( y_\alpha \) be the unique element of \( A_{-\alpha} \) such that \([x_\alpha, y_\alpha]_A = h_\alpha\). Thus \( y_\alpha \neq 0 \), because \( h_\alpha \neq 0 \), so that \( A_{-\alpha} \) consists exactly of multiples of \( y_\alpha \) by elements of \( k \).

Remember that \( A((-\Phi^+)) \) is the linear subspace of \( A \) spanned by \( A_{-\alpha} \), as in Sections 22.4 and 31.1. More precisely, \( A((-\Phi^+)) \) is a Lie subalgebra of \( A \), as in Section 22.6. Because \( A((-\Phi^+)) \) and \( B_\Delta \) are Lie subalgebras of \( A \), there are natural injective homomorphisms from universal enveloping algebras \( UA((-\Phi^+)) \) and \( UB_\Delta \) of \( A((-\Phi^+)) \) and \( B_\Delta \), respectively, into \( UA \), as associative algebras over \( k \), as in Section 25.12. Thus we may identify \( UA((-\Phi^+)) \) and \( UB_\Delta \) with subalgebras of \( UA \) in this way.

By construction, \( A \) corresponds to the direct sum of \( A((-\Phi^+)) \) and \( B_\Delta \), as a vector space over \( k \). This leads to an isomorphism between \( UA((-\Phi^+)) \otimes (UB_\Delta) \), as vector spaces over \( k \), as in Section 25.12. In particular,

\[
(31.8.2) \quad \text{every element of } UA \text{ can be expressed as a finite sum, where each term corresponds to the product of an element of } UA((-\Phi^+) \text{ and an element of } UB_\Delta.
\]

Observe that

\[
(31.8.3) \quad (UB_\Delta) \cdot v = \{tv : t \in k\},
\]

because \( v \) is an eigenvector of \( B_\Delta \), by hypothesis. This also uses the fact that \( UB_\Delta \) is generated as an associative algebra over \( k \) by its multiplicative identity element and the image of \( B_\Delta \) in \( UB_\Delta \). It follows that

\[
(31.8.4) \quad V = (UA((-\Phi^+))) \cdot v,
\]

by (31.8.1) and (31.8.2).

Let \( \beta_1, \ldots, \beta_r \) be a list of the elements of \( \Phi^+ \), so that \( y_{\beta_1}, \ldots, y_{\beta_r} \) is a basis for \( A((-\Phi^+)) \), as a vector space over \( k \). If \( m_1, \ldots, m_r \) are nonnegative integers, then

\[
(31.8.5) \quad y_{\beta_1}^{m_1} \cdots y_{\beta_r}^{m_r}
\]

corresponds to an element of \( UA((-\Phi^+)) \subseteq UA \). More precisely, \( y_{\beta_j} \) is identified with an element of \( UA((-\Phi^+)) \) for each \( j = 1, \ldots, r \), so that \( y_{\beta_j}^{m_j} \) corresponds to an element of \( UA((-\Phi^+)) \) as well. Of course, this is interpreted as being the multiplicative identity element when \( m_j = 0 \). Thus the product of \( y_{\beta_j}^{m_j} \) over \( j = 1, \ldots, r \) corresponds to an element of \( UA((-\Phi^+)) \) too.
The elements of $UA(-\Phi^+)$ of the form (31.8.5), where $m_j$ is a nonnegative integer for each $j = 1, \ldots, r$, form a basis for $UA(-\Phi^+)$ as a vector space over $k$, by the Poincaré–Birkhoff–Witt theorem. In particular,

$$\text{(31.8.6)} \quad UA(-\Phi^+) \text{ is spanned by elements of the form (31.8.5), as a vector space over } k.$$

If $m_j$ is a nonnegative integer for each $j = 1, \ldots, r$, then (31.8.5) acts on $v$ to get an element

$$\text{(31.8.7)} \quad (y_{j_1}^{m_1} \cdots y_{j_r}^{m_r}) \cdot v$$

of $V$. In fact,

$$\text{(31.8.8)} \quad V \text{ is spanned by elements of the form (31.8.7), as a vector space over } k.$$

This follows from (31.8.4) and (31.8.6). This corresponds to the first part of part (a) of the theorem on p108 of [13], part (1) of Proposition 2 on p57 of [23], and part of the proof of Theorem 3.1 on p58 of [24].

If $m_j$ is a nonnegative integer for each $j = 1, \ldots, r$, then (31.8.7) has weight

$$\text{(31.8.9)} \quad \sum_{j=1}^r m_j j$$

as an element of $V$, in the sense of Section 31.5. This follows from (31.5.3), because $v \in V_\mu$, by hypothesis. This implies that (31.8.7) has weight of the form

$$\text{(31.8.10)} \quad \mu = \sum_{\alpha \in \Delta} c_\alpha \alpha,$$

where $c_\alpha$ is a nonnegative integer for each $\alpha \in \Delta$, because $\beta_j \in \Phi^+$ for every $j = 1, \ldots, r$.

In particular, $V$ is spanned by its weight spaces, by (31.8.8). This means that

$$\text{(31.8.11)} \quad V \text{ corresponds to the direct sum of its weight spaces, as a vector space over } k,$$

as in Section 31.5. We also get that

$$\text{(31.8.12)} \quad \text{all of the weights of } V \text{ are of the form mentioned in (31.8.9), and thus (31.8.10).}$$

This corresponds to the second part of part (a) and part (b) of the theorem on p108 of [13], the first part of part (2) of Proposition 2 on p57 of [23], and part (b) of Theorem 3.1 on p58 of [24].

Remember that $\Delta$ is supposed to be a basis for $A_0$, as a vector space over $k$, as in Section 31.1. In particular, the elements of $\Delta$ are linearly independent.
in $A'_0$, as a vector space over $k$. If $c_\alpha$ is a nonnegative integer for each $\alpha \in \Delta$, then there are only finitely many nonnegative integers $m_1, \ldots, m_r$ such that

$$\sum_{j=1}^{r} m_j \beta_j = \sum_{\alpha \in \Delta} c_\alpha \alpha.$$  

(31.8.13)

This uses the facts that $\beta_j \in \Phi^+$ for every $j = 1, \ldots, r$, and $k$ has characteristic $0$, by hypothesis. Similarly, if $c_\alpha = 0$ for every $\alpha \in \Delta$, and $m_1, \ldots, m_r$ are nonnegative integers, then (31.8.13) holds only when $m_j = 0$ for every $j = 1, \ldots, r$.

It follows that

$$\text{the weight spaces in } V \text{ have finite dimension},$$

as in the first part of part (c) of the theorem on p108 of [13], and the second part of part (2) of Proposition 2 on p57 of [23]. We also get that

$$\dim V_\mu = 1,$$

(31.8.15)

as in the second part of part (c) of the theorem on p108 of [13], part (3) of Proposition 2 on p57 of [23], and part (c) of Theorem 3.1 on p58 of [24]. Note that $\mu$ may be called the highest weight of $V$, because all of the weights of $V$ are of the form (31.8.10), with $c_\alpha \geq 0$ for every $\alpha \in \Delta$.

### 31.9 Indecomposability

Let us continue with the same notation and hypotheses as in the previous section. We would like to show that $V$ is indecomposable as a module over $A$, which is to say that it cannot be expressed is the direct sum of two nontrivial submodules of $V$. This corresponds to the first part of part (c) of the theorem on p108 of [13], and to part (4) of Proposition 2 on p57 of [23].

Suppose that $V$ corresponds to the direct sum of two linear subspaces $Y$ and $Z$, and that $Y$ and $Z$ are submodules of $V$, as a module over $A$. Thus we get linear projection mappings from $V$ onto $Y$ and $Z$ whose kernels are $Z$ and $Y$, respectively, and these projections commute with the action of $A$ on $V$. If $\nu \in A'_0$, then one can use this to check that these projection mappings send $V_\nu$ into $Y_\nu$ and $Z_\nu$, respectively, where $V_\nu$, $Y_\nu$ and $Z_\nu$ are as in Section 31.5. This means that $V_\nu$ corresponds to the direct sum of $Y_\nu$ and $Z_\nu$, as a vector space over $k$.

In particular, if we take $\nu = \mu$, then we get that one of $Y_\mu$ or $Z_\mu$ has dimension one, by (31.8.15). This implies that $v$ is an element of $Y$ or $Z$, so that $Y$ or $Z$ is all of $V$, because $V$ is generated by $v$ as a module over $A$, as in (31.8.1). Of course, this means that the other of $Y$ or $Z$ is $\{0\}$, so that $V$ is indecomposable, as desired. This is the argument near the top of p58 of [23].

If $V$ has finite dimension as a vector space over $k$, then it follows that $V$ is irreducible as a module over $A$, by Weyl’s theorem. This uses the hypothesis
that $k$ have characteristic 0, as well as the semisimplicity of $A$, as in Section 31.1. This corresponds to part (d) of Proposition 3 on p60 of [23], and part (a) of Theorem 3.1 on p58 of [24].

Let $Z$ be any submodule of $V$ as a module over $A$, so that $Z$ is a linear subspace of $V$, and the action of $A$ on $V$ maps $Z$ into itself. Part (d) of the theorem on p108 of [13] states that

$$Z \text{ corresponds to the direct sum of its weight spaces,}$$

as a vector space over $k$.

Of course, it suffices to check that $Z$ is spanned by its weight spaces, as in Section 31.5.

Suppose for the sake of a contradiction that there is an element of $Z$ that is not in the linear span of the weight spaces of $Z$. Because $V$ is spanned by its weight spaces, every element of $Z$ can be expressed as

$$z = \sum_{j=1}^{n} z_j,$$

where for each $j = 1, \ldots, n$, $z_j \in V_{\nu_j}$ for some $\nu_j \in A'_0$. Let $n$ be the smallest positive integer for which there is a $z \in Z$ that is not in the linear span of the weight spaces of $Z$, and which can be expressed as in (31.9.2). Of course, this means that $n \geq 2$, and that the $\nu_j$’s are distinct elements of $A'_0$. We also have that $z_j \notin Z$ for each $j = 1, \ldots, n$.

Because $\nu_1 \neq \nu_2$, there is a $w \in A_0$ such that $\nu_1(w) \neq \nu_2(w)$. Observe that

$$w \cdot z = \sum_{j=1}^{n} w \cdot z_j = \sum_{j=1}^{n} \nu_j(w) z_j,$$

so that

$$w \cdot z - \nu_1(w) z = \sum_{j=2}^{n} (\nu_j(w) - \nu_1(w)) z_j.$$

The left side is an element of $Z$, and the minimality of $n$ implies that it is in the linear span of the weight spaces of $Z$. Remember that the linear span of the weight spaces in $V$ corresponds to the direct sum of the weight spaces in $V$, as a vector space over $k$, as in Section 31.5. It follows that the terms in the sum on the right side of (31.9.4) are elements of $Z$. This implies that $z_2 \in Z$, because $\nu_2(w) \neq \nu_1(w)$, by construction. This is a contradiction, as in the preceding paragraph.

Let $Z$ be a submodule of $V$ as a module over $A$ again, and suppose that $Z \neq V$. This implies that $v \notin Z$, by (31.8.1), and thus

$$Z \cap V_{\mu} = \{0\},$$

by (31.8.15). Let $V^0$ be the linear subspace of $V$ spanned by the weight spaces $V_{\nu}$ with $\nu \in A'_0$ and $\nu \neq \mu$. This is a proper linear subspace of $V$, because $V$
31.10. IRREDUCIBLE STANDARD CYCLIC MODULES

corresponds to the direct sum of its weight spaces, and \( V_\mu \neq \{0\} \), by hypothesis. Using (31.9.5), we get that

\[
Z \subseteq V^0,
\]

because \( Z \) is spanned by its weight spaces, as before.

Let \( V^1 \) be the linear subspace of \( V \) spanned by all of the proper submodules of \( V \), as a module over \( A \). Thus

\[
V^1 \subseteq V^0,
\]

by (31.9.6). In particular, this implies that \( V \) is indecomposable as a module over \( A \), as on p109 of [13]. More precisely, \( V \) is not spanned by proper submodules, as a module over \( A \).

It is easy to see that \( V^1 \) is a submodule of \( V \), as a module over \( A \). In fact, \( V^1 \) is the unique maximal proper submodule of \( V \), as a module over \( A \), as in part (e) of the theorem on p108 of [13]. The quotient \( V/V^1 \) is a module over \( A \) in a natural way, because \( V^1 \) is a submodule of \( V \). Observe that

\[
V/V^1 \text{ is irreducible, as a module over } A,
\]

as in part (e) of the theorem on p108 of [13], because \( V^1 \) is a maximal proper submodule of \( V \). Of course, \( V/V^1 \neq \{0\} \), because \( V^1 \neq V \).

If the quotient of \( V \) by a proper submodule is irreducible as a module over \( A \), then that submodule of \( V \) is maximal. This means that \( V/V^1 \) is the only nontrivial irreducible quotient of \( V \), as a module over \( A \), as in part (e) of the theorem on p108 of [13].

31.10 Irreducible standard cyclic modules

Let us use the same notation and hypotheses as in Section 31.1 again, and let \( V \) be a module over \( A \), as a Lie algebra over \( k \). Suppose that \( v \in V \) is a primitive or maximal vector of weight \( \mu \in \mathcal{A}_0' \), and that \( V \) is irreducible as a module over \( A \). Let \( U_A \) be a universal enveloping algebra of \( A \), so that \( (U_A) \cdot v \) is a submodule of \( V \) that contains \( v \). Because \( V \) is irreducible, we get that \( V = (U_A) \cdot v \), so that \( V \) is standard cyclic of weight \( \mu \). This corresponds to the first step in the proof of Theorem 1 on p58 of [23].

Thus \( V \) satisfies the hypotheses in Section 31.8, and so has the same properties as before. This corresponds to part (b) of Theorem 1 on p58 of [23].

Suppose that \( v_1 \in V \) is another primitive or maximal vector, with weight \( \mu_1 \in \mathcal{A}_0' \). It follows that

\[
V = (U_A) \cdot v_1,
\]

because \( V \) is irreducible, as before. This means that the analogues of the statements in Section 31.8 for \( v_1 \) and \( \mu_1 \) hold as well.

Because \( \mu_1 \) is a weight of \( V \), there are nonnegative integers \( c_\alpha, \alpha \in \Delta \), such that

\[
\mu_1 = \mu - \sum_{\alpha \in \Delta} c_\alpha \alpha,
\]

(31.10.2)
as in (31.8.12). Similarly, there are nonnegative integers $c_{1,\alpha}, \alpha \in \Delta$, such that
\begin{equation}
\mu = \mu_1 - \sum_{\alpha \in \Delta} c_{1,\alpha} \alpha,
\end{equation}
by the analogue of (31.8.12) for $\mu_1$, and the fact that $\mu$ is a weight of $V$. It follows that $c_{1,\alpha} = -c_\alpha$ for every $\alpha \in \Delta$, because $\Delta$ is a basis for $A_0$, and $k$ has characteristic 0, as in Section 31.1. This implies that $c_\alpha = c_{1,\alpha} = 0$ for every $\alpha \in \Delta$, so that
\begin{equation}
\mu_1 = \mu.
\end{equation}
We also get that
\begin{equation}
v_1 = t_1 v \text{ for some } t_1 \in k \setminus \{0\},
\end{equation}
by (31.8.15).

This corresponds to the corollary on p109 of [13], part (a) of Theorem 1 on p58 of [13], and part (1) of Theorem 3.2 on p59 of [24]. Note that $\mu$ may be called the highest weight of $V$, as in [23, 24]. This terminology may also be used for any standard cyclic module of weight $\mu$, as in [13]. Of course, this refers to the expression of other weights as in (31.8.10).

Let $Z$ be another irreducible module over $A$ with a primitive or maximal vector $z$. If $Z$ is isomorphic to $V$ as a module over $A$, then $Z$ has highest weight $\mu$ too, because of (31.10.4). Conversely, suppose that $Z$ has highest weight $\mu$, and let us show that $Z$ is isomorphic to $V$, as modules over $A$. This corresponds to Theorem A on p109 of [13], part (c) of Theorem 1 on p58 of [24], and part (2) of Theorem 3.2 on p59 of [24].

Let $Y$ be the direct sum of $V$ and $Z$, which is the same as the Cartesian product of $V$ and $Z$, where addition and scalar multiplication are defined coordinatewise. More precisely, $Y$ is a module over $A$, where the action of $A$ is defined coordinatewise as well. Observe that
\begin{equation}
y = (v, z)
\end{equation}
is a primitive or maximal vector of weight $\mu$ in $Y$. Put
\begin{equation}
E = (UA) \cdot y,
\end{equation}
which is the submodule of $Y$, as a module over $A$, generated by $y$. Thus $E$ is a standard cyclic module of weight $\mu$, by construction.

Let $p_V, p_Z$ be the obvious projections from $Y$ onto $V, Z$, respectively, and let $\phi_V, \phi_Z$ be their restrictions to $E$. Of course, these mappings are module homomorphisms, and
\begin{equation}
\phi_V(y) = v, \phi_Z(y) = z.
\end{equation}
This implies that
\begin{equation}
\phi_V(E) = V, \phi_Z(E) = Z,
\end{equation}
because $V, Z$ are standard cyclic with respect to $v, z$, respectively, as before. The kernels of $\phi_V, \phi_Z$ are
\begin{equation}
\ker \phi_V = (\{0\} \times Z) \cap E, \ker \phi_Z = (V \times \{0\}) \cap E.
\end{equation}
Of course, $V \times \{0\}$ and $\{0\} \times Z$ may be identified with $V$ and $Z$, respectively, as modules over $A$.

Note that the only elements of $E$ of weight $\mu$ are multiples of $y$ by elements of $k$, as in (31.8.15). Thus there are no nonzero elements of the kernels of $\phi_V$ or $\phi_Z$ of weight $\mu$. It follows that $(0,z) \notin \ker \phi_V$, $(v,0) \notin \ker \phi_Z$, so that

$$\ker \phi_V \neq \{0\} \times Z, \quad \ker \phi_Z \neq V \times \{0\}. \quad (31.10.11)$$

We can identify $\ker \phi_V$, $\ker \phi_Z$ with submodules of $Z$, $V$, respectively, as modules over $A$. These submodules are proper, by (31.10.11).

This implies that these submodules are trivial, because $V$ and $Z$ are irreducible, by hypothesis. This means that the kernels of $\phi_V$ and $\phi_Z$ are trivial, so that $\phi_V$, $\phi_Z$ are isomorphisms from $E$ onto $V$, $Z$, respectively, as modules over $A$. Of course, this leads to an isomorphism from $V$ onto $Z$. This is the argument on p58f of [23].

The argument on p109 of [13] uses the fact that $V$ and $Z$ are isomorphic to irreducible quotients of $E$, while there is only one irreducible quotient of $E$, as in Section 31.9.

If $V$ and $Z$ have finite dimension as vector spaces over $k$, then $E$ has finite dimension too. This implies that $E$ is irreducible as a module over $A$, because $E$ is indecomposable, as in the previous section. In this case, one can get that the kernels of $\phi_V$ and $\phi_Z$ are trivial because they are proper submodules of $E$. This is the argument used on p59 of [24], where the modules are taken to be finite-dimensional.

### 31.11 Representations and left ideals

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $A$. Also let $UA$ be a universal enveloping algebra of $A$ with multiplicative identity element $e = e_{UA}$ and associated mapping $i = i_{UA}$ from $A$ into $UA$, as in Section 25.4.

Suppose that $\mathcal{I}$ is a left ideal in $UA$, as an associative algebra over $k$. The quotient $(UA)/\mathcal{I}$ can be defined as a module over $k$, and in fact as a left module over $UA$. Let $q_{\mathcal{I}}$ be the natural quotient mapping from $UA$ onto $(UA)/\mathcal{I}$, so that $q_{\mathcal{I}}(e)$ is an element of the quotient. Of course,

$$\text{(31.11.1)} \quad (UA) \cdot q_{\mathcal{I}}(e) = (UA)/\mathcal{I},$$

by construction. Note that $(UA)/\mathcal{I}$ may be considered as a module over $A$ too, as a Lie algebra over $k$.

Suppose that $V$ is a module over $k$ that is a module over $A$ as a Lie algebra over $k$. Thus $V$ may be considered as a (left) module over $UA$, as an associative algebra over $UA$, as before. If $v \in V$, then

$$\text{(31.11.2)} \quad \mathcal{I}_v = \{x \in UA : x \cdot v = 0\}$$
is a left ideal in \( UA \). This means that the quotient \( (UA)/I_v \) is defined as a left module over \( UA \). The mapping
\[
x \mapsto x \cdot v
\]
from \( UA \) into \( V \) leads to an injective homomorphism from \( (UA)/I_v \) into \( V \), as left modules over \( UA \). More precisely, this mapping sends \( (UA)/I_v \) onto \( (UA) \cdot v \), which is a submodule of \( V \), as a module over \( UA \). If \( V = (UA) \cdot v \), then we get an isomorphism from \( (UA)/I_v \) onto \( V \), as left modules over \( UA \).

### 31.12 Tensor products over algebras

Let \( k \) be a commutative ring with a multiplicative identity element, and let \( \mathcal{A} \) be an associative algebra over \( k \), with a multiplicative identity element \( e = e_\mathcal{A} \). Also let \( V, W \) be modules over \( k \), where \( V \) is a right module over \( \mathcal{A} \), and \( W \) is a left module over \( \mathcal{A} \). The tensor product of \( V \) and \( W \) over \( \mathcal{A} \) is a module \( V \otimes_\mathcal{A} W \) over \( k \) with the following two properties.

First, the tensor product is equipped with a mapping from \( V \otimes W \) into \( V \otimes_\mathcal{A} W \) that is bilinear over \( k \). If \( v \in V \) and \( w \in W \), then the image of \((v, w)\) under this mapping is denoted \( v \odot w \), as usual. This mapping should also satisfy
\[
(v \cdot a) \odot w = v \odot (a \cdot w)
\]
for every \( a \in \mathcal{A}, v \in V, \) and \( w \in W \).

Second, let \( Z \) be any module over \( k \), and let \( b \) be a mapping from \( V \times W \) into \( Z \) that is bilinear over \( k \). Under these conditions, there should be a unique homomorphism \( c \) from \( V \otimes_\mathcal{A} W \) into \( Z \), as modules over \( k \), such that
\[
b(v, w) = c(v \odot w)
\]
for every \( v \in V \) and \( w \in W \).

The tensor product can be obtained using a standard construction, and it is unique up to a suitable isomorphic equivalence. Note that \( V \otimes_\mathcal{A} W \) is generated as a module over \( k \) by elements of the form \( v \odot w, v \in V, w \in W \).

Let \( I \) be a nonempty set, and let \( V_j \) be a right module over \( \mathcal{A} \) for every \( j \in I \). Thus \( \bigoplus_{j \in I} V_j \) is a right module over \( \mathcal{A} \) too, and one can check that
\[
\left( \bigoplus_{j \in I} V_j \right) \otimes_\mathcal{A} W = \bigoplus_{j \in I} (V_j \otimes_\mathcal{A} W).
\]
Similarly, if \( W_j \) is a left module over \( \mathcal{A} \) for every \( j \in I \), then \( \bigoplus_{j \in I} W_j \) is a left module over \( \mathcal{A} \), and one can verify that
\[
V \otimes_\mathcal{A} \left( \bigoplus_{j \in I} W_j \right) = \bigoplus_{j \in I} (V \otimes_\mathcal{A} W_j).
\]
31.12. TENSOR PRODUCTS OVER ALGEBRAS

Let $V_1, V_2$ be right modules over $\mathcal{A}$, and let $W_1, W_2$ be left modules over $\mathcal{A}$. Suppose that $\phi_{V_1}$ is a homomorphism from $V_1$ into $V_2$, as right modules over $\mathcal{A}$, and that $\phi_{W_1}$ is a homomorphism from $W_1$ into $W_2$, as left modules over $\mathcal{A}$. Consider the mapping from $V_1 \times W_1$ into $V_2 \otimes_{\mathcal{A}} W_2$ defined by

\begin{equation}
(v_1, w_1) \mapsto \phi_{V_2}(v_1) \otimes \phi_{W_1}(w_1).
\end{equation}

This mapping is bilinear over $k$, and satisfies

\begin{equation}
\phi_{V_2}(a \cdot v_1) \otimes \phi_{W_1}(w_1) = (\phi_{V_2}(v_1) \cdot a) \otimes \phi_{W_1}(w_1)
\end{equation}

for every $a \in \mathcal{A}$, $v_1 \in V_1$, and $w_1 \in W_1$. This leads to a unique homomorphism $\phi$ from $V_1 \otimes_{\mathcal{A}} W_1$ into $V_2 \otimes_{\mathcal{A}} W_2$, as modules over $k$, such that

\begin{equation}
\phi(v_1 \otimes w_1) = \phi_{V_2}(v_1) \otimes \phi_{W_1}(w_1)
\end{equation}

for every $v_1 \in V_1$ and $w_1 \in W_1$.

Let $V_2$ be another right module over $\mathcal{A}$, and let $W_3$ be another left module over $\mathcal{A}$. Suppose that $\psi_{V_2}$ is a homomorphism from $V_2$ into $V_3$, as right modules over $\mathcal{A}$, and that $\psi_{W_3}$ is a homomorphism from $W_2$ into $W_3$, as left modules over $\mathcal{A}$. This leads to a homomorphism $\psi$ from $V_2 \otimes_{\mathcal{A}} W_2$ into $V_3 \otimes_{\mathcal{A}} W_3$, as modules over $k$, as in the previous paragraph. Of course, $\psi_{V_3} \circ \phi_{V_1}$ is a homomorphism from $V_1$ into $V_3$, as right modules over $\mathcal{A}$, and $\psi_{W_3} \circ \phi_{W_1}$ is a homomorphism from $W_1$ into $W_3$, as left modules over $\mathcal{A}$. One can verify that $\psi \circ \phi$ is the same as the homomorphism from $V_1 \otimes_{\mathcal{A}} W_1$ into $V_3 \otimes_{\mathcal{A}} W_3$, as modules over $k$, obtained from $\psi_{V_3} \circ \phi_{V_1}, \psi_{W_3} \circ \phi_{W_1}$ as before.

Suppose that $\mathcal{A}_1$ is an associative algebra over $k$ that contains $\mathcal{A}$ as a subalgebra, and with the same multiplicative identity element $e$. We may consider $\mathcal{A}_1$ as a left and a right module over itself, and over $\mathcal{A}$. Let us consider $\mathcal{A}_1$ as a right module over $\mathcal{A}$ for the moment, and let $W$ be a left module over $\mathcal{A}$ again. Thus $\mathcal{A}_1 \otimes_{\mathcal{A}} W$ can be defined as a module over $k$, as before.

If $a_1 \in \mathcal{A}_1$, then

\begin{equation}x \mapsto a_1 x
\end{equation}

defines a mapping from $\mathcal{A}_1$ into itself that is linear over $k$. More precisely, (31.12.9) is a homomorphism from $\mathcal{A}_1$ into itself, as a right module over itself, and in particular as a right module over $\mathcal{A}$. Using this and the identity mapping on $W$, we get a homomorphism from $\mathcal{A}_1 \otimes W$ into itself, as a module over $k$. One can check that $\mathcal{A}_1 \otimes_{\mathcal{A}} W$ becomes a left module over $\mathcal{A}$ in this way.

In particular, we can take $\mathcal{A}_1 = \mathcal{A}$ here, so that $\mathcal{A} \otimes_{\mathcal{A}} W$ may be considered as a left module over $\mathcal{A}$. Of course, we already have a mapping from $\mathcal{A} \times W$ into $W$ that is bilinear over $k$, given by the action of $\mathcal{A}$ on $W$ as a left module over $\mathcal{A}$. It is easy to see that $W$ satisfies the requirements of the tensor product $\mathcal{A} \otimes_{\mathcal{A}} W$, as a module over $k$, with respect to this mapping from $\mathcal{A} \times W$ into $W$. One can verify that the action of $\mathcal{A}$ on $\mathcal{A} \otimes_{\mathcal{A}} W$ on the left just mentioned corresponds exactly to the given action of $\mathcal{A}$ on $W$ on the left.
31.13 Existence of standard cyclic modules

We continue to use the same notation and hypotheses as in Section 31.1. Let \( \mu \in A_0^* \) be given. We would like to show that there is a standard cyclic module over \( A \) with weight \( \mu \).

Remember that \( \Delta \) is a base for the root system \( \Phi \), and that \( B_\Delta \) is the standard Borel subalgebra of \( A \) associated to \( A_0 \) and \( \Delta \). We also have that \( B_\Delta \) corresponds to the direct sum of \( A_0 \) and \( N_\Delta = A(\Phi^+) \) as a vector space over \( k \), and that \( [B_\Delta, B_\Delta] = N_\Delta \), as before.

Let \( L(\mu) \) be a one-dimensional vector space over \( k \). We can define an action of \( B_\Delta \) on \( L(\mu) \) by putting

\[
(w + x) \cdot u = \mu(w) u
\]

for every \( w \in A_0 \), \( x \in N_\Delta \), and \( u \in L(\mu) \). It is easy to see that this makes \( L(\mu) \) into a module over \( B_\Delta \), as a Lie algebra over \( k \).

Let \( UA \) be a universal enveloping algebra of \( A \), with multiplicative identity element \( e = e_{UA} \) and associated mapping \( i = i_{UA} \) from \( A \) into \( UA \), as in Section 25.4. We may as well take the universal enveloping algebra \( UB_\Delta \) of \( B_\Delta \) to be the subalgebra of \( UA \) generated by \( e \) and \( i(B_\Delta) \), as in Section 25.12. We may also consider \( L(\mu) \) as a left module over \( UB_\Delta \), as an associative algebra over \( k \). Remember that \( i \) is injective, as in Section 25.12, so that we may as well identify \( A \) with its image in \( UA \).

Of course, \( UA \) may be considered as a right module over itself, and thus a right module over \( UB_\Delta \). Let us take

\[
Z(\mu) = (UA) \otimes_{UB_\Delta} L(\mu),
\]

where the right side is as in the previous section. More precisely, this is a module over \( k \), and in fact a left module over \( UA \). This means that \( Z(\mu) \) may be considered as a module over \( A \), as a Lie algebra over \( k \).

Let \( u_0 \) be a nonzero element of \( L(\mu) \), and put

\[
v = e \otimes u_0 \in Z(\mu).
\]

One can check that

\[
(UA) \cdot v = Z(\mu),
\]

because \( L(\mu) \) has dimension one, as a vector space over \( k \). If \( w \in A_0 \), then

\[
w \cdot v = w \cdot (e \otimes u_0) = w \cdot u_0 = e \otimes (w \cdot u_0) = \mu(w) (e \otimes u_0).
\]

Similarly, if \( x \in N_\Delta \), then

\[
x \cdot v = x \cdot (e \otimes u_0) = x \otimes u_0 = e \otimes (x \cdot u_0) = 0.
\]

Remember that \( A(-\Phi^+) \) is a Lie subalgebra of \( A \), and that \( A \) corresponds to the direct sum of \( A(-\Phi^+) \) and \( B_\Delta \), as linear subspaces of \( A \). This leads to
an isomorphism between $UA$ and $(UA(-\Phi^+)) \otimes (UB_\Delta)$, as vector spaces over $k$, as in Section 25.12. One can use this to get that $UA$ is free as a right module over $UB_\Delta$. This corresponds to part of Corollary D on p92 of [13], aside from using right modules here.

More precisely, let $x_\alpha \in A_\alpha$ and $y_{\alpha} \in A_{-\alpha}$ be as in Section 31.8 for each $\alpha$ in $\Phi^+$, and let $\beta_1, \ldots, \beta_r$ be a list of the elements of $\Phi^+$ again. Thus $y_{\beta_1}, \ldots, y_{\beta_r}$ form a basis for $A(-\Phi^+)$, as a vector space over $k$, and

$$(31.13.7) \quad \{y_{\beta_1}^{m_1} \cdots y_{\beta_r}^{m_r} : m_1, \ldots, m_r \in \mathbb{Z}_+ \cup \{0\}\}$$

is a basis for $UA(-\Phi^+)$, as a vector space over $k$, by the Poincaré–Birkhoff–Witt theorem, as before. One can use the Poincaré–Birkhoff–Witt theorem again to get that $UA$ is free as a right module over $UB_\Delta$, with basis (31.13.7).

Indeed, the Poincaré–Birkhoff–Witt theorem implies that $UA$ corresponds to the direct sum of the linear subspaces

$$(31.13.8) \quad y_{\beta_1}^{m_1} \cdots y_{\beta_r}^{m_r} (UB_\Delta)$$

as a vector space over $k$, where $m_1, \ldots, m_r$ are nonnegative integers. The Poincaré–Birkhoff–Witt theorem also ensures that

$$(31.13.9) \quad z \mapsto y_{\beta_1}^{m_1} \cdots y_{\beta_r}^{m_r} z$$

is a one-to-one linear mapping from $UB_\Delta$ onto (31.13.8) for all $m_1, \ldots, m_r \geq 0$. This can be seen by combining $y_{\beta_1}, \ldots, y_{\beta_r}$ with a basis for $B_\Delta$, as a vector space over $k$, to get a basis for $A$, as a vector space over $k$.

Clearly (31.13.8) is a submodule of $UA$, as a right module over $UB_\Delta$, for every $m_1, \ldots, m_r \geq 0$. In fact, (31.13.9) is an isomorphism from $UB_\Delta$ onto (31.13.8), as right modules over $UB_\Delta$, for every $m_1, \ldots, m_r \geq 0$. This means that $UA$ corresponds to the direct sum of the submodules (31.13.8) as a right module over $UB_\Delta$, and that $UA$ is free as a right module over $UB_\Delta$.

It follows that $Z(\mu)$ corresponds to the direct sum of

$$(31.13.10) \quad (y_{\beta_1}^{m_1} \cdots y_{\beta_r}^{m_r} (UB_\Delta)) \otimes L(\mu)$$

as a vector space over $k$, where $m_1, \ldots, m_r$ are nonnegative integers. Note that (31.13.10) is isomorphic to $L(\mu)$ as a vector space over $k$ for every $m_1, \ldots, m_r \geq 0$, as in the previous section. This means that (31.13.10) is a one-dimensional vector space over $k$ spanned by

$$(31.13.11) \quad (y_{\beta_1}^{m_1} \cdots y_{\beta_r}^{m_r}) \cdot (e \otimes u_0) = (y_{\beta_1}^{m_1} \cdots y_{\beta_r}^{m_r}) \cdot v$$

for every $m_1, \ldots, m_r \geq 0$, where $v$ is as in (31.13.3).

In particular, (31.13.11) is nonzero for all nonnegative integers $m_1, \ldots, m_r$. It follows that $v \neq 0$, by taking $m_1 = \cdots = m_r = 0$. This shows that $v$ is a primitive or maximal vector in $Z(\mu)$ with weight $\mu$. Thus $Z(\mu)$ is standard cyclic of weight $\mu$, as a module over $A$. We also get that the elements (31.13.11)
of $Z(\mu)$, with $m_1, \ldots, m_r \geq 0$, form a basis for $Z(\mu)$, as a vector space over $k$. This corresponds to some of the remarks on p110 of [13], and on p59 of [23]. This corresponds to part of Exercise 6 on p62 of [24] as well.

Observe that

\[(31.13.12) \quad z \mapsto z \cdot (e \otimes u_0)\]

is a one-to-one linear mapping from $UA(-\Phi^+)$ onto $Z(\mu)$, because (31.13.7) is a basis for $UA(-\Phi^+)$, as a vector space over $k$. Note that $Z(\mu)$ may be considered as a left module over $UA(-\Phi^+)$, where $UA(-\Phi^+)$ is identified with a subalgebra of $UA$ in the usual way. In fact, (31.13.12) is an isomorphism from $UA(-\Phi^+)$ onto $Z(\mu)$, as left modules over $UA(-\Phi^+)$, as mentioned on p110 of [13].

As in Section 31.9, $Z(\mu)$ has a unique nontrivial irreducible quotient, as a module over $A$. It is easy to see that the quotient is also standard cyclic of weight $\mu$, using the image of $v$ in the quotient. This corresponds to Theorem B on p110 of [13], Theorem 2 on p59 of [23], and part of Exercise 6 on p62 of [24].
Chapter 32

Representations and semisimplicity, 2

32.1 Weights and finite dimension

As before, we use the same notation and hypotheses as in Section 31.1. In particular, \((A, [\cdot, \cdot]_A)\) is a semisimple Lie algebra of positive finite dimension over a field \(k\) of characteristic 0, and \(A_0\) is a Lie subalgebra of \(A\) consisting of ad-diagonalizable elements of \(A\). We also let \(\Delta\) be a base for the root system \(\Phi\), and \(B_\Delta\) be the standard Borel subalgebra of \(A\) associated to \(A_0\) and \(\Delta\), as before.

Let \(V\) be a vector space over \(k\) that is a module over \(A\), as a Lie algebra over \(k\). If \(\mu \in A_0\), then \(V_\mu\) is the linear subspace of \(V\) consisting of \(v \in V\) such that \(w \cdot v = \mu(w) v\) for every \(w \in A_0\), as in Section 31.5.

In this section, we suppose also that \(k\) is algebraically closed, and that \(V\) has finite dimension, as a vector space over \(k\). Under these conditions, it is well known that

\[
(32.1.1) \quad V \text{ corresponds to the direct sum of } V_\mu \text{ for finitely many } \mu \in A_0, \text{ as a vector space over } k.
\]

This is mentioned on p107 of [13], and corresponds to part (a) of Proposition 3 on p60 of [23], and to Proposition 2.2 on p57 of [24].

To see this, note that for each \(w \in A_0\),

\[
(32.1.2) \quad v \mapsto w \cdot v
\]

is diagonalizable on \(V\), as in Section 14.9. These linear mappings on \(V\) commute with each other, because \(A_0\) is commutative as a Lie subalgebra of \(A\), and \(V\) is a module over \(A\). This implies (32.1.1), which is the same as saying that the linear mappings (32.1.2) with \(w \in A_0\) can be simultaneously diagonalized on \(V\). This is the argument indicated in [13, 23].
Alternatively, let $V$ be the linear subspace of $V$ spanned by the $V_\mu$’s, $\mu \in A_0^\prime$, as in Section 31.5. Remember that $\bar{V}$ corresponds to the direct sum of the $V_\mu$’s, and that $\bar{V}$ is a submodule of $V$, as a module over $A$, as before. If $\bar{V} \neq V$, then Weyl’s theorem implies that $V$ corresponds to the direct sum of $\bar{V}$ and a nonzero submodule $Y$, as a module over $A$.

If $w \in A_0$, then (32.1.2) maps $Y$ into itself, because $Y$ is a submodule of $V$, as a module over $A$. This implies that (32.1.2) has a nonzero eigenvector in $Y$, because $k$ is algebraically closed, and $Y$ is a vector space over $k$ of positive finite dimension. It follows that there is a nonzero element $y$ of $Y$ that is a simultaneous eigenvector of (32.1.2) for every $w \in A_0$, because these linear mappings commute on $V$ and thus $Y$, as before. This means that $y \in V_\mu$ for some $\mu \in A_0^\prime$, so that $y \in \bar{V}$, which is a contradiction. This is the argument used in the proof of Proposition 2.2 on p57 of [24].

If $V \neq \{0\}$, then $V$ has a primitive or maximal vector, in the sense of Section 31.6. This is mentioned on p108 of [13], and corresponds to part (c) of Proposition 3 on p60 of [23], and to Proposition 2.7 on p58 of [24]. To see this, remember that $B_\Delta$ is solvable as a Lie algebra over $k$, so that there is a nonzero element $v$ of $V$ that is an eigenvector of the action of every element of $B_\Delta$ on $V$, by Lie’s theorem. This implies that $v$ is a primitive or maximal vector, as before.

Another argument is mentioned on p58 of [24], as follows. Remember that for each $\alpha \in \Phi$, there is a unique element $h_\alpha$ of $[A_\alpha, A_{-\alpha}]$ such that $\alpha(h_\alpha) = 2 = 1 + 1$ in $k$, as in Section 31.1. Note that $h_{-\alpha} = -h_\alpha$ for every $\alpha \in \Phi$. If $\alpha \in \Phi$ and $\nu \in A_0^\prime$ is a weight of $V$, then

\begin{equation}
\nu(h_\alpha) \in \mathbb{Z},
\end{equation}

using the natural embedding of $\mathbb{Q}$ into $k$. This is part (b) of Proposition 3 on p60 of [23], which corresponds to a remark after the statement of the theorem on p112 of [13]. It suffices to verify this when $\alpha$ is an element of the set $\Phi^+$ of positive roots with respect to $\Delta$.

Let $\alpha \in \Phi^+$ be given, let $x_\alpha$ be a nonzero element of $A_\alpha$, and let $y_\alpha$ be an element of $A_{-\alpha}$ such that $[x_\alpha, y_\alpha]_A = h_\alpha$. The linear span of $x_\alpha$, $y_\alpha$, and $h_\alpha$ in $A$ is a Lie subalgebra of $A$ that is isomorphic to $sl_2(k)$. We may consider $V$ as a module over this subalgebra of $A$, as a Lie algebra over $k$. The eigenvalues of the action of $h_\alpha$ on $V$ correspond to integers, as in Section 15.9. These eigenvalues are the same as the values of the weights of $V$ at $h_\alpha$, as desired.

The condition (32.1.3) is also mentioned in Exercise 2 on p62 of [24], and could be obtained from Theorem 3.1 on p58 and Theorem 4.1 on p59 of [24].
32.2 Weights and a larger field

Let us continue with the same notation and hypotheses as in the previous section, except that we no longer ask $k$ to be algebraically closed. Let $k_1$ be an algebraically closed field with $k \subseteq k_1$.

Remember that we can get a vector space $\tilde{V}$ over $k_1$ from $V$ as in Section 31.2. In particular, the dimension of $\tilde{V}$ as a vector space over $k_1$ is the same as the dimension of $V$ as a vector space over $k$, which is finite, by hypothesis. As before, we may consider $V$ as a linear subspace of $\tilde{V}$, as a vector space over $k$.

If $T$ is a linear mapping from $V$ into itself, then $T$ may be considered as a linear mapping from $V$ into $\tilde{V}$, as a vector space over $k$. This has a unique extension to a linear mapping $\tilde{T}$ from $\tilde{V}$ into itself, as a vector space over $k_1$.

Suppose for the moment that $\tilde{T}$ is diagonalizable on $\tilde{V}$, with distinct eigenvalues $\lambda_1, \ldots, \lambda_l \in k_1$. This implies that
\[
\prod_{j=1}^{l}(\tilde{T} - \lambda_j \tilde{I}_V) = 0
\]
(32.2.1)
as a linear mapping on $\tilde{V}$, where $\tilde{I}_V$ is the identity mapping on $\tilde{V}$. If
\[
(32.2.2)
\lambda_1, \ldots, \lambda_l \in k,
\]
then it follows that
\[
\prod_{j=1}^{l}(T - \lambda_j I_V) = 0
\]
on $V$, where $I_V$ is the identity mapping on $V$. This means that $T$ is diagonalizable on $V$, by standard results.

We also get a Lie algebra $(\tilde{A}, [\cdot, \cdot], \cdot)$ over $k_1$ from $A$, as in Section 31.2. The action of $A$ on $V$ has a unique extension to an action of $\tilde{A}$ on $\tilde{V}$, that makes $\tilde{V}$ a module over $\tilde{A}$, as a Lie algebra over $k_1$, as before.

The conditions on $A$ discussed in Section 31.1 lead to analogous properties of $\tilde{A}$, as in Section 31.4. In particular, $\tilde{A}$ is semisimple as a Lie algebra over $k_1$, which could also be obtained from the semisimplicity of $A$ as in Section 11.5.

Let $(\tilde{A}_0)$ be the linear subspace of $\tilde{A}$ corresponding to $A_0$, which is a Lie subalgebra of $\tilde{A}$ that is commutative as a Lie algebra over $k_1$, as in Section 31.4. Remember that the elements of $(A_0)$ are ad-diagonalizable in $\tilde{A}$.

Let $(\tilde{A}_0)'$ be the dual of $(\tilde{A}_0)$, as a vector space over $k_1$, as before. If $\mu \in (\tilde{A}_0)'$, then we take $\tilde{V}_\mu$ to be the set of $v \in \tilde{V}$ such that $w \cdot v = \mu(w) v$ for every $w \in (\tilde{A}_0)'$, as in Section 31.5. Because $k_1$ is algebraically closed, we have that $\tilde{V}$ corresponds to the direct sum of $\tilde{V}_\mu$ for finitely many $\mu \in (\tilde{A}_0)'$, as in the previous section.

If $\alpha \in A_0'$, then let $\tilde{\alpha}$ be the unique extension of $\alpha$ to an element of $(\tilde{A}_0)'$, as in Section 31.4. Remember that the analogue $\Phi$ of $\Phi$ for $\tilde{A}$ and $(\tilde{A}_0)$ consists of $\tilde{\alpha}$, $\alpha \in \Phi$. 

If $\alpha \in \Phi$, then $h_\alpha$ is the unique element of $[A_\alpha, A_{-\alpha}]$ with $\alpha(h_\alpha) = 2$, as in Section 31.1. We may also consider $h_\alpha$ as an element of $\tilde{A}$, and indeed it is the unique element of $[\tilde{A}_\alpha, \tilde{A}_{-\alpha}] = [(\tilde{A}_\alpha), (\tilde{A}_{-\alpha})]$ with $\tilde{\alpha}(h_\alpha) = 2$.

If $\alpha \in \Phi$ and $\nu \in (A_0)'$ is a weight of $\tilde{V}$, then $\nu(h_\alpha)$ corresponds to an integer, as in the previous section. This means that the action of $h_\alpha$ on $\tilde{V}$ is diagonalizable, and that the eigenvalues of this action correspond to integers, and thus elements of $k$.

The action of $h_\alpha$ on $\tilde{V}$ is the same as the unique extension of the action of $h_\alpha$ on $V$ to a linear mapping from $\tilde{V}$ into itself, by construction. It follows that the action of $h_\alpha$ on $V$ is diagonalizable too, as before.

Remember that $A_0$ is spanned by the $h_\alpha$'s, $\alpha \in \Phi$, as in Section 31.1. This means that the action of any element of $A_0$ on $V$ is diagonalizable, because the action of $h_\alpha$ is diagonalizable for every $\alpha \in \Phi$.

The actions of the elements of $A_0$ on $V$ commute with each other, because $A_0$ is commutative as a Lie subalgebra of $A$, and $V$ is a module over $A$, as a Lie algebra over $k$. This implies that the actions of the elements of $A_0$ on $V$ can be simultaneously diagonalized, as in the previous section.

This shows that (32.1.1) holds, even if $k$ is not necessarily algebraically closed. This corresponds to part of Exercise 2 on p62 of [24].

If $V \neq \{0\}$, then one can check that $V$ has a primitive or maximal vector, using the statement in the preceding paragraph, and the second argument for the analogous assertion in the previous section. This corresponds to another part of Exercise 2 on p62 of [24].

If $V \neq \{0\}$ and $V$ is irreducible as a module over $A$, then it follows that $V$ is standard cyclic as a module over $A$, as in the remarks at the beginning of Section 31.10. This corresponds to the corollary to Proposition 3 on p60 of [23], and is mentioned in the proof of part (1) of Theorem 3.2 on p59 of [24]. This also corresponds to some remarks on p108f of [13].

### 32.3 Conditions for finite dimension

We continue to use the same notation and hypotheses as in Section 31.1. Let $V$ be a vector space over $k$ that is a module over $A$, as a Lie algebra over $k$. In this section, we suppose that $V$ is irreducible as a module over $A$, and that $V$ is standard cyclic of weight $\mu \in A_0'$.

Suppose for the moment that $V$ has finite dimension, as a vector space over $k$. We would like to show that

\begin{equation}
(32.3.1) \quad \text{for every } \alpha \in \Phi^+, \mu(h_\alpha) \text{ is a nonnegative integer.}
\end{equation}

This corresponds to the theorem on p112 of [13], to part of Theorem 3 on p60 of [23], and part of Theorem 4.1 on p59 of [24].

Let $\alpha \in \Phi^+$ be given. Remember that $\mu(h_\alpha)$ corresponds to an integer, using the natural embedding of $\mathbb{Q}$ into $k$, as in (32.1.3). This also uses the remarks in the previous section, when $k$ is not necessarily algebraically closed.
As before, let \( x_\alpha \) be a nonzero element of \( A_\alpha \), and let \( y_\alpha \) be an element of \( A_{-\alpha} \) such that \([x_\alpha, y_\alpha]_A = h_\alpha\), so that the linear span of \( x_\alpha \), \( y_\alpha \), and \( h_\alpha \) in \( A \) is a Lie subalgebra of \( A \) isomorphic to \( sl_2(k) \). Thus \( V \) may be considered as a module over \( sl_2(k) \) as well.

By hypothesis, there is a \( v \in V \) that is primitive or maximal of weight \( \mu \) in \( V \), as a module over \( A \). This implies that \( v \) is primitive or maximal of weight \( \mu(h_\alpha) \) in \( V \) as a module over the linear span of \( x_\alpha \), \( y_\alpha \), and \( h_\alpha \), as in Section 15.2. It follows that \( \mu(h_\alpha) \geq 0 \), as in Section 15.3.

Conversely, it is well known that (32.3.1) implies that the dimension of \( V \) is finite under these conditions. This corresponds to part of the theorem on p113 of [13], to part of Theorem 3 on p60 of [23], and part of Theorem 4.1 on p59 of [24] when \( A = sl_n(k) \). Before we begin the proof of this, let us review some facts related to inverse roots.

Remember from Section 31.1 that \( \Phi \) is a reduced root system in the finite-dimensional vector space \( E_\mathbb{R} \) over the real numbers. If \( \alpha \in \Phi \), then \( \lambda_\alpha \) is the linear functional on \( E_\mathbb{R} \) associated to the symmetry on \( E_\mathbb{R} \) with vector \( \alpha \) that maps \( \Phi \) onto itself in the usual way. We have seen that the collection \( \Phi' \) of \( \lambda_\alpha \), \( \alpha \in \Phi \), is a root system in the dual \( E_\mathbb{R}' \) of \( E_\mathbb{R} \), as in Section 19.8. Remember that \( \Delta \) is a base for \( \Phi \), and let \( \Delta' \) be the set of \( \lambda_\alpha \), \( \alpha \in \Delta \). We have also seen that \( \Delta' \) is a base for \( \Phi' \), because \( \Phi \) is reduced, as in Section 19.13.

Using \( \Delta' \), we can define \( (\Phi')^+ = (\Phi')^{+}\Delta' \) to be the set of dual roots in \( \Phi' \) that are positive with respect to \( \Delta' \), which is to say that they can be expressed as linear combinations of the elements of \( \Delta' \) with nonnegative coefficients. Equivalently, \( (\Phi')^+ \) consists exactly of \( \lambda_\alpha \), \( \alpha \in \Phi^+ \), as in Section 30.3.

If \( \alpha \in \Phi \), then
\[
(32.3.2) \quad \gamma(h_\alpha) = \lambda_\alpha(\gamma)
\]
for every \( \gamma \in \Phi \), as in (31.1.2). More precisely, the right side is an integer, by the definition of a root system, which corresponds to an element of \( k \) under the natural embedding of \( \mathbb{Q} \) into \( k \). Remember that \( h_\alpha \) is uniquely determined by (32.3.2), because \( A_\alpha' \) is spanned by \( \Phi \), as in Section 31.1.

If \( \alpha \in \Phi \), then \( \lambda_\alpha \) can be expressed as a linear combination of \( \lambda_\beta \), \( \beta \in \Delta \), with integer coefficients, because \( \Delta' \) is a base for \( \Phi' \). This implies that \( h_\alpha \) can be expressed as a linear combination of \( h_\beta \), \( \beta \in \Delta \), with the same integer coefficients, because \( h_\alpha \) is uniquely determined by (32.3.2). Similarly, if \( \alpha \in \Phi^+ \), then \( \lambda_\alpha \) can be expressed as a linear combination of \( \lambda_\beta \), \( \beta \in \Delta \), with nonnegative integer coefficients, because \( \lambda_\alpha \in \Phi^+ \), as before. This means that \( h_\alpha \) can be expressed as a linear combination of \( h_\beta \), \( \beta \in \Delta \), with the same nonnegative integer coefficients.

It follows that (32.3.1) holds when \( \mu(h_\beta) \) is a nonnegative integer for every \( \beta \in \Delta \). This corresponds to a remark just after Theorem 3 on p60 of [23]. Similarly, if \( \nu \in A_\alpha' \), then \( \nu(h_\alpha) \) corresponds to an integer for every \( \alpha \in \Phi \) if and only if \( \nu(h_\beta) \) corresponds to an integer for every \( \beta \in \Delta \).
CHAPTER 32. REPRESENTATIONS AND SEMISIMPLICITY, 2

32.4 Connections with abstract weights

We continue to use the same notation and hypotheses as in Section 31.1, as usual. Let \( \Phi' \) and \( \Delta' \) be as in the previous section. Remember that \( E_Q \) is the linear span of \( \Phi \) in \( A'_0 \), as a vector space over \( Q \). We may also consider \( E_Q \) to be the linear span of \( \Phi \) in \( E_R \), as a vector space over \( Q \).

If \( \alpha \in \Phi \), then \( \lambda_\alpha \) takes integer values on \( \Phi \), by the definition of a root system. This implies that \( \lambda_\alpha \) maps \( E_Q \) into \( Q \), so that the restriction of \( \lambda_\alpha \) to \( E_Q \) may be considered as a linear functional on \( E_Q \), as a vector space over \( Q \).

If \( z \in E_R \), then \( \lambda_\alpha(z) \) is uniquely determined by \( \lambda_\alpha(z) \), \( \alpha \in \Delta \), because the \( \lambda_\alpha \)'s, \( \alpha \in \Delta \), form a basis for \( E_R \). If \( \lambda_\alpha(z) \in Q \) for every \( \alpha \in \Delta \), then there is an element of \( E_Q \) for which the values of \( \lambda_\alpha \), \( \alpha \in \Delta \), are the same, as in the preceding paragraph. This means that \( z \in E_Q \), because \( z \) is determined by \( \lambda_\alpha(z), \alpha \in \Delta \). Thus

\[
(32.4.1) \quad E_Q = \{ z \in E_R : \lambda_\alpha(z) \in Q \text{ for every } \alpha \in \Delta \},
\]

because \( \lambda_\alpha \) maps \( E_Q \) into \( Q \) for every \( \alpha \in \Phi \), as before.

If \( \gamma \in \Phi \), then \( \gamma(h_\alpha) \) corresponds to an integer, using the natural embedding of \( Q \) into \( k \), by (32.3.2). Similarly, if \( \gamma \in E_Q \), then \( \gamma(h_\alpha) \) corresponds to an element of \( Q \) as in (32.3.2). Every function of \( \alpha \in \Delta \) with values in \( Q \) can be obtained as \( \gamma(h_\alpha) \) for some \( \gamma \in E_Q \), as before.

Remember that \( h_\alpha \), \( \alpha \in \Delta \), is a basis for \( A_0 \) as a vector space over \( k \), as in Section 31.1. In particular, every element of \( A'_0 \) is uniquely determined by its values on the \( h_\alpha \)'s, \( \alpha \in \Delta \). If \( \nu \in A'_0 \) and \( \nu(h_\alpha) \) corresponds to an element of \( Q \) for every \( \alpha \in \Delta \), under the natural embedding of \( Q \) into \( k \), then it follows that \( \nu \in E_Q \). This shows that

\[
(32.4.2) \quad E_Q = \{ \nu \in A'_0 : \nu(h_\alpha) \in Q \text{ for every } \alpha \in \Delta \}.
\]

Put

\[
(32.4.3) \quad \Upsilon = \Upsilon_\Phi = \{ z \in E_R : \lambda_\alpha(z) \in Z \text{ for every } \alpha \in \Phi \},
\]

as in Section 30.1. Equivalently,

\[
(32.4.4) \quad \Upsilon = \{ z \in E_R : \lambda_\alpha(z) \in Z \text{ for every } \alpha \in \Delta \},
\]
32.5. SOME INITIAL STEPS

because the collection $\Delta'$ of $\lambda_\alpha$, $\alpha \in \Delta$, is a base for $\Phi'$, as before. In particular,

$$(32.4.5) \quad \mathcal{Y} \subseteq E_Q.$$

by (32.4.1). Thus

$$(32.4.6) \quad \mathcal{Y} = \{ z \in E_Q : \lambda_\alpha(z) \in \mathbb{Z} \text{ for every } \alpha \in \Delta \} = \{ z \in E_Q : \lambda_\alpha(z) \in \mathbb{Z} \text{ for every } \alpha \in \Phi \}.$$

Let us check that

$$(32.4.7) \quad \mathcal{Y} = \{ \nu \in A'_0 : \nu(h_\alpha) \in \mathbb{Z} \text{ for every } \alpha \in \Delta \}.$$

If $\nu \in A'_0$ satisfies $\nu(h_\alpha) \in \mathbb{Z}$ for every $\alpha \in \Delta$, then $\nu \in E_Q$, by (32.4.2). In this case,

$$(32.4.8) \quad \nu(h_\alpha) = \lambda_\alpha(\nu)$$

for every $\alpha \in \Phi$, by (32.3.2). This implies that $\nu \in \mathcal{Y}$, by (32.4.6). We also have that $\mathcal{Y}$ is contained in the right side of (32.4.7), by (32.4.6) and (32.4.8), and because $E_Q \subseteq A'_0$, by construction. Similarly,

$$(32.4.9) \quad \mathcal{Y} = \{ \nu \in A'_0 : \nu(h_\alpha) \in \mathbb{Z} \text{ for every } \alpha \in \Phi \}.$$

This can be obtained from (32.4.6), or the fact that if $\alpha \in \Phi$, then $h_\alpha$ can be expressed as a linear combination of $h_\beta$, $\beta \in \Delta$, with integer coefficients, as in the previous section.

As in Section 30.3, we put

$$(32.4.10) \quad \mathcal{Y}^+ = \mathcal{Y}^+_{\Phi,\Delta} = \{ z \in \mathcal{Y} : \lambda_\alpha(z) \geq 0 \text{ for every } \alpha \in \Delta \}.$$

This is the same as

$$(32.4.11) \quad \mathcal{Y}^+ = \{ z \in \mathcal{Y} : \lambda_\alpha(z) \geq 0 \text{ for every } \alpha \in \Phi^+ \},$$

where $\Phi^+$ is the set of positive roots with respect to $\Delta$, as usual. This is a bit simpler here, because $\Phi$ is reduced as a root system. Using (32.4.5) and (32.4.8), we get that

$$(32.4.12) \quad \mathcal{Y}^+ = \{ \nu \in \mathcal{Y} : \nu(h_\alpha) \geq 0 \text{ for every } \alpha \in \Delta \} = \{ \nu \in \mathcal{Y} : \nu(h_\alpha) \geq 0 \text{ for every } \alpha \in \Phi^+ \}.$$

An element of $A'_0$ may be called integral if it is in $\mathcal{Y}$, and dominant integral if it is in $\mathcal{Y}^+$, as on p.112 of [13].

32.5 Some initial steps

We continue to use the same notation and hypotheses as in Section 31.1. If $\alpha$ is in the base $\Delta$ for $\Phi$, then let $x_\alpha$ be a nonzero element of $A_\alpha$, and let $y_\alpha$ be
an element of $A_\alpha$ such that $[x_\alpha, y_\alpha]_A = h_\alpha$, as before. If $\alpha, \beta \in \Delta$ and $\alpha \neq \beta$, then $[x_\alpha, y_\beta]_A = 0$, as in Section 23.5. Indeed, it is easy to see that $\alpha - \beta \notin \Phi$ in this case, so that $A_{\alpha - \beta} = \{0\}$.

Let $UA$ be a universal enveloping algebra of $A$ again, and let us identify $A$ with its image in $UA$ under the associated Lie algebra homomorphism, as before. If $\alpha, \beta \in \Delta$ and $\alpha \neq \beta$, then $[x_\alpha; y_\beta]_A = 0$, as in Section 23.5. Indeed, it is easy to see that $\alpha \neq 2 \Delta$ in this case, so that $A_\alpha = f_0 g_0$.

Let $UA$ be a universal enveloping algebra of $A$ again, and let us identify $A$ with its image in $UA$ under the associated Lie algebra homomorphism, as before. If $\alpha, \beta \in \Delta$ and $\alpha \neq \beta$, then $[x_\alpha; y_\beta] = 0$ for every positive integer $l$. More precisely, $y_\beta^l$ is defined as an element of $UA$, and the left side uses the commutator bracket associated to multiplication in $UA$. Of course, (32.5.1) holds when $l = 1$ as in the preceding paragraph, which implies the analogous statement for $l \geq 2$. This corresponds to part (a) of the lemma on p113 of [13].

If $\alpha, \beta \in \Delta$, then

$$[x_\alpha, y_\beta^l] = 0$$

for every positive integer $l$. More precisely, $y_\beta^l$ is defined as an element of $UA$, and the left side uses the commutator bracket associated to multiplication in $UA$. Of course, (32.5.1) holds when $l = 1$ as in the preceding paragraph, which implies the analogous statement for $l \geq 2$. This corresponds to part (a) of the lemma on p113 of [13].

If $\alpha, \beta \in \Delta$, then

$$[x_\alpha, y_\beta^l] = -l \beta(h_\alpha) y_\beta^l$$

for every positive integer $l$. If $l = 1$, then this reduces to the fact that $[x_\alpha, y_\beta]_A = -\beta(h_\alpha)$, because $h_\alpha \in A_0$ and $y_\beta \in A_\beta$. Otherwise, one can use induction and the fact that the commutator with $h_\alpha$ defines a derivation on $UA$, as in Section 2.5. This corresponds to part (b) of the lemma on p113 of [13].

If $\alpha \in \Delta$, then

$$[x_\alpha, y_\alpha^l] = -l y_\alpha^{l-1} ((l-1) \cdot e - h_\alpha)$$

for every positive integer $l$, where $e = e_{UA}$ is the multiplicative identity element in $UA$. If $l = 1$, then $y_\alpha^{l-1}$ is interpreted as being equal to $e$, and (32.5.3) follows from the way that $y_\alpha$ was chosen. Otherwise, if $l \geq 2$, then

$$[x_\alpha, y_\alpha^l] = [x_\alpha, y_\alpha] y_\alpha^{l-1} + y_\alpha [x_\alpha, y_\alpha^{l-1}] = h_\alpha y_\alpha^{l-1} + y_\alpha [x_\alpha, y_\alpha^{l-1}],$$

because the commutator with $x_\alpha$ defines a derivation on $UA$, as before. Using induction, we may suppose that the analogue of (32.5.3) for $l - 1$ holds, to get that

$$[x_\alpha, y_\alpha^l] = h_\alpha y_\alpha^{l-1} - (l-1) y_\alpha^{l-1} ((l-2) \cdot e - h_\alpha).$$

Observe that

$$h_\alpha y_\alpha^{l-1} = [h_\alpha, y_\alpha^{l-1}] + y_\alpha^{l-1} h_\alpha = -2 (l-1) y_\alpha^{l-1} + y_\alpha^{l-1} h_\alpha,$$

by (32.5.2). One can combine this with (32.5.5) to get (32.5.3), as desired. This corresponds to part (c) of the lemma on p113 of [13].

Let $V$ be a vector space over $k$ that is a module over $A$, as a Lie algebra over $k$. Suppose that $V$ is irreducible as a module over $A$, that $V$ is standard cyclic of weight $\mu \in A'_0$, and that

$$m_\alpha = \mu(h_\alpha)$$

is a nonnegative integer for every $\alpha \in \Delta$. We would like to show that $V$ has finite dimension, as in Section 32.3.
32.5. SOME INITIAL STEPS

By hypothesis, there is a primitive or maximal vector \( v \in V \) of weight \( \mu \) such that \( V = (UA) \cdot v \). Let \( \alpha \in \Delta \) be given, and put

\[
v_{\alpha} = y_{\alpha}^{m_{\alpha}+1} \cdot v.
\]

(32.5.8)

If \( \beta \in \Delta \) and \( \alpha \neq \beta \), then

\[
x_{\beta} \cdot v_{\alpha} = y_{\alpha}^{m_{\alpha}+1} \cdot (x_{\beta} \cdot v) = 0.
\]

(32.5.9)

This uses (32.5.1) in the first step, and the fact that \( v \) is primitive or maximal in the second step.

Observe that

\[
x_{\alpha} \cdot v_{\alpha} = ([x_{\alpha}, y_{\alpha}^{m_{\alpha}+1}]) \cdot v + y_{\alpha}^{m_{\alpha}+1} \cdot (x_{\alpha} \cdot v) = -(m_{\alpha} + 1) y_{\alpha}^{m_{\alpha}} \cdot ((m_{\alpha} \cdot e - h_{\alpha}) \cdot v),
\]

using (32.5.3) and the fact that \( x_{\alpha} \cdot v = 0 \) in the second step. However,

\[
(m_{\alpha} \cdot e - h_{\alpha}) \cdot v = m_{\alpha} v - h_{\alpha} \cdot v = m_{\alpha} v - \mu(h_{\alpha}) v = 0,
\]

because \( v \) has weight \( \mu \) in \( V \). Thus

\[
x_{\alpha} \cdot v_{\alpha} = 0.
\]

(32.5.12)

Note that \( v_{\alpha} \) has weight \( \mu - (m_{\alpha} + 1)\alpha \) with respect to \( A_{0} \), as in Section 31.5. If \( v_{\alpha} \neq 0 \), then \( v_{\alpha} \) would be a primitive or maximal vector in \( V \), because of (32.5.9) and (32.5.12). This is not possible, because \( V \) is irreducible as a module over \( A \) and \( m_{\alpha} \neq -1 \), as in Section 31.10. This shows that

\[
y_{\alpha}^{m_{\alpha}+1} \cdot v = v_{\alpha} = 0,
\]

(32.5.13)
as in step (1) on p113 of [13].

Alternatively, the linear span of \( x_{\alpha}, y_{\alpha}, \) and \( h_{\alpha} \) in \( A \) is a Lie subalgebra of \( A \) that is isomorphic to \( sl_{2}(k) \). Thus \( V \) may be considered as a module over this subalgebra, and \( v \) is primitive or maximal of weight \( m_{\alpha} \) in the sense of Section 15.2. One can use this to get (32.5.12) as in Section 15.3. This is the argument used on p61 of [23].

Using (32.5.13), we get that

\[
\text{the linear subspace of } V \text{ spanned by } y_{\alpha}^{j} \cdot v, \ j = 0, \ldots, m_{\alpha},
\]

is a submodule of \( V \), as a module over the Lie subalgebra of \( A \) spanned by \( x_{\alpha}, y_{\alpha}, \) and \( h_{\alpha} \).

This follows from the remarks in Section 15.3, as on p61 of [23]. Alternatively, this linear subspace of \( V \) is mapped into itself by the action of \( y_{\alpha} \), because of (32.5.13). This linear subspace of \( V \) is also mapped into itself by the action of \( h_{\alpha} \), because \( y_{\alpha}^{j} \cdot v \) has weight \( \mu - j \alpha \) in \( V \) for each \( j \geq 0 \), as in Section 31.5. One can check that this linear subspace of \( V \) is mapped into itself by the action of \( x_{\alpha} \), using (32.5.3) and the fact that \( x_{\alpha} \cdot v = 0 \). This is the argument used in step (2) on p113 of [13]. Note that this linear subspace of \( V \) is nontrivial, because \( v \neq 0 \).
32.6 Submodules over subalgebras

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$. Suppose that $V$ is a module over $k$ that is a module over $A$, as a Lie algebra over $k$. Let $C$ be a Lie subalgebra of $A$, so that $V$ may be considered as a module over $C$ as well. Also let $Z$ be a submodule of $V$, as a module over $k$, that is a submodule of $V$, as a module over $C$ as well.

Consider the subset $A \cdot Z$ of $V$ consisting of finite sums of elements of the form $a \cdot z$, where $a \in A$ and $z \in Z$. This is clearly a submodule of $V$, as a module over $k$. Let us check that $A \cdot Z$ is a submodule of $V$ as a module over $C$. This corresponds to part of the argument in step (3) on p113 of [13], and to a remark on p61 of [23].

If $a \in A$, $c \in C$, and $z \in Z$, then

$$c \cdot (a \cdot z) = ([c, a)_A] \cdot z + a \cdot (c \cdot z).$$

(32.6.1)

Of course, $([c, a]_A) \cdot z \in A \cdot Z$ automatically, and $c \cdot z \in Z$, because $Z$ is a submodule of $V$, as a module over $C$. It follows that $c \cdot (a \cdot z) \in A \cdot Z$, which implies that $A \cdot Z$ is a submodule of $V$, as a module over $C$.

Suppose now that $k$ is a field, and that $A$ has finite dimension, as a vector space over $k$. Let $V(C)$ be the linear subspace of $V$ spanned by all finite-dimensional linear subspaces $Z$ of $V$ such that $Z$ is a submodule of $V$, as a module over $C$. We would like to verify that $V(C)$ is a submodule of $V$, as a module over $A$. This corresponds to another part of the argument in step (3) on p113 of [13], and to another remark on p61 of [23].

Let $Z$ be a finite-dimensional linear subspace of $V$ that is a submodule of $V$, as a module over $C$. Thus $A \cdot Z$ is a submodule of $V$, as a module over $C$, as before. It is easy to see that $A \cdot Z$ has finite dimension as a vector space over $k$, because $A$ and $Z$ have finite dimension. This implies that

$$A \cdot Z \subseteq V(C).$$

(32.6.2)

One can use this to check that $V(C)$ is a submodule of $V$, as a module over $A$.

If $Z_1, \ldots, Z_n$ are finitely many finite-dimensional linear subspaces of $V$ that are submodules of $Z_1, \ldots, Z_n$ as a module over $C$, then it is easy to see that the linear subspace of $V$ spanned by $Z_1, \ldots, Z_n$ has finite dimension and is a submodule of $V$, as a module over $C$. This implies that every element of $V(C)$ is contained in a finite-dimensional linear subspace of $V$ that is a submodule of $V$, as a module over $C$. Similarly, every finite-dimensional linear subspace of $V(C)$ is contained in a finite-dimensional linear subspace of $V$ that is a submodule of $V$, as a module over $C$.

Let us return now to the same notation and hypotheses as in Section 31.1, so that $A$ is a finite-dimensional Lie algebra over a field $k$ of characteristic 0 in particular. Let $V$ be a vector space over $k$ that is a module over $A$ again, as a Lie algebra over $k$, and suppose that $V$ is irreducible as a module over $A$, that $V$ is standard cyclic of weight $\mu \in A_0$, and that $\mu(h_\alpha)$ corresponds to a nonnegative integer for each $\alpha \in \Delta$, with respect to the standard embedding.
of $Q$ into $k$, as in the previous section. If $\alpha \in \Delta$, then let $x_\alpha$ and $y_\alpha$ be as before, and let $C_\alpha$ be the linear span of $x_\alpha$, $y_\alpha$, and $h_\alpha$ in $A$. Thus $C_\alpha$ is a Lie subalgebra of $A$ that is isomorphic to $sl_2(k)$.

Under these conditions, there is a nontrivial finite-dimensional linear subspace of $V$ that is a submodule of $V$, as a module over $C_\alpha$, as in the previous section. This means that $V(C_\alpha) \neq \{0\}$, where $V(C_\alpha)$ is as defined earlier. It follows that

$$V(C_\alpha) = V,$$

because $V(C_\alpha)$ is a submodule of $V$, as a module over $A$, as before. This corresponds to step (3) on p113 of [13], and to an argument on p61 of [23].

### 32.7 Some related exponentials on $V$

Let us continue with the same notation and hypotheses as in the previous section, including those from Section 31.1. Thus $V$ is a vector space over $k$ that is a module over $A$, as a Lie algebra over $k$, $V$ is irreducible as a module over $A$, $V$ is standard cyclic of weight $2\Delta$, with respect to the standard embedding of $Q$ into $k$, as before. If $\alpha \in \Delta$, then we let $x_\alpha, y_\alpha \in A$ be as before, so that the linear span $C_\alpha$ of $x_\alpha$, $y_\alpha$, and $h_\alpha$ in $A$ is a Lie subalgebra of $A$ that is isomorphic to $sl_2(k)$.

Let $\rho$ be the representation of $A$ on $V$, so that $\rho(a) = a \cdot v$ for every $a \in A$ and $v \in V$. If $Z$ is a finite-dimensional linear subspace of $V$ that is a submodule of $V$, as a module over $C_\alpha$, then $\rho_{x_\alpha}$ and $\rho_{y_\alpha}$ are nilpotent on $Z$. Indeed, $Z$ is isomorphic as a module over $C_\alpha$ to a direct sum of finitely many irreducible finite-dimensional modules over $C_\alpha$, by Weyl’s theorem. Every irreducible finite-dimensional module over $C_\alpha$ is isomorphic to the module $W(m)$ described in Section 15.4 for some nonnegative integer $m$, as in Sections 15.5 and 15.9. It follows that $\rho_{x_\alpha}$ and $\rho_{y_\alpha}$ are nilpotent on $Z$, because of the analogous statement for $W(m)$, as in Section 15.8. This implies that $\rho_{x_\alpha}$ and $\rho_{y_\alpha}$ are locally nilpotent on $V$, as in Section 27.12, because of (32.6.3). This corresponds to step (4) on p114 of [13].

Of course, $\rho_{-y_\alpha} = -\rho_{y_\alpha}$ is locally nilpotent on $V$ as well. Thus $\exp \rho_{x_\alpha}$ and $\exp \rho_{-y_\alpha}$ are defined as linear mappings from $V$ into itself, as in Section 27.12. More precisely, these are one-to-one linear mappings from $V$ onto itself, as before. It follows that

$$\theta_\alpha = (\exp \rho_{x_\alpha}) \circ (\exp \rho_{-y_\alpha}) \circ (\exp \rho_{x_\alpha}),$$

defines a one-to-one linear mapping from $V$ onto itself, as in step (5) on p114 of [13].

Let us check that

$$\rho_{h_\alpha} \circ \theta_\alpha = -\theta_\alpha \circ \rho_{h_\alpha}$$

on $V$. It suffices to verify that this holds on any finite-dimensional linear subspace $Z$ of $V$ that is a submodule of $V$, as a module over $C_\alpha$, because of (32.6.3). Remember that $Z$ is isomorphic as a module over $C_\alpha$ to the direct sum of finitely
many irreducible finite-dimensional modules over $C_{\alpha}$, each of which is isomorphic to a module of the form $W(m)$ as in Section 15.4. Thus (32.7.2) follows from the analogous statement for $W(m)$, as in Section 15.8.

If $h \in A_0$ and $\alpha(h) = 0$, then $[h, x_{\alpha}]_A = [h, y_{\alpha}]_A = 0$. This means that $\rho_h$ commutes with $\rho_{x_{\alpha}}$ and $\rho_{y_{\alpha}}$ on $V$, and thus with $\exp \rho_{x_{\alpha}}$, $\exp \rho_{y_{\alpha}}$. It follows that

$$\rho_h \circ \theta_\alpha = \theta_\alpha \circ \rho_h$$

(32.7.3)
on $V$ under these conditions.

Let $\nu \in A'_0$ be given, and let $V_\nu$ be the linear subspace of $V$ consisting of vectors with weight $\nu$ with respect to $A_0$, as in Section 31.5. Thus $V_\nu$ consists of $z \in V$ such that $\rho_h(z) = \nu(h) z$ for every $h \in A_0$. More precisely, it suffices to check this when $h = h_{\alpha}$, and when $\alpha(h) = 0$, because $A_0$ is spanned by $h_{\alpha}$ and the kernel of $\alpha$, as a vector space over $k$. This uses the fact that $\alpha(h_{\alpha}) = 2 = 1 + 1 \neq 0$ in $k$, because $k$ has characteristic 0, by hypothesis.

Put

$$\tau_\alpha(\nu) = \nu - \nu(h_{\alpha}) \alpha,$$

which is also an element of $A'_0$. Note that $\tau_\alpha(\nu) = \nu$ on the kernel of $\alpha$, that $(\tau_\alpha(\nu))(h_{\alpha}) = -\nu(h_{\alpha})$, and that $\tau_\alpha(\nu)$ is uniquely determined in $A'_0$ by these conditions. One can check that

$$\theta_\alpha(V_\nu) = V_{\tau_\alpha(\nu)},$$

(32.7.4)using (32.7.2) and (32.7.3). This corresponds to step (6) on p114 of [13], and to part of Remark (1) on p62 of [23].

It follows that

$$\dim V_\nu = \dim V_{\tau_\alpha(\nu)},$$

(32.7.5)as vector spaces over $k$. In particular, $V_\nu \neq \{0\}$ exactly when $V_{\tau_\alpha(\nu)} \neq \{0\}$. This is the same as saying that $\nu$ is a weight of $V$ with respect to $A_0$ if and only if $\tau_\alpha(\nu)$ is a weight of $V$.

Alternatively, suppose that $V_\nu \neq \{0\}$, and let $z$ be a nonzero element of $V_\nu$. Because $V$ is standard cyclic of weight $\mu$ as a module over $A$, $\nu$ can be expressed as $\mu$ minus a linear combination of roots with integer coefficients, as in Section 31.8. In particular, this implies that $\nu(h_{\alpha})$ corresponds to an integer, under the standard embedding of $Q$ into $k$, because $\mu(h_{\alpha})$ corresponds to an integer, by hypothesis.

Let $Z$ be a finite-dimensional linear subspace of $V$ that is a submodule of $V$, as a module over $C_{\alpha}$, such that $z \in Z$, as in the previous section. Put

$$u = (\rho_{h_{\alpha}})^{\nu(h_{\alpha})}(z) \quad \text{when } \nu(h_{\alpha}) \geq 0,$$

$$u = (\rho_{x_{\alpha}})^{-\nu(h_{\alpha})}(z) \quad \text{when } \nu(h_{\alpha}) \leq 0,$$

(32.7.7)which is an element of $Z$. One can verify that $u \neq 0$, because $Z$ is isomorphic as a module over $C_{\alpha}$ to the direct sum of finitely many modules over $C_{\alpha}$ of the form $W(m)$ as in Section 15.4, as before.
We also have that \( u \) has weight \((32.7.4)\) in \( V \) with respect to \( A_0 \), as in Section 31.5. This implies that \( V_{\tau_\alpha(\nu)} \neq \{0\} \), as before. Conversely, if \( V_{\tau_\alpha(\nu)} \neq \{0\} \), then the same argument shows that \( \nu \neq \{0\} \), because \( \tau_\alpha(\tau_\alpha(\nu)) = \nu \). This corresponds to an argument on p61 of [23].

### 32.8 Weights and the Weyl group

We continue to use the same notation and hypotheses as in Section 31.1. Remember that \( \Phi \) is a root system in \( E_\mathbb{R} \), which is a vector space over the real numbers of positive finite dimension, and that \( E_\mathbb{Q} \) is the same as the linear span of \( \Phi \) in \( E_\mathbb{R} \) as a vector space over \( \mathbb{Q} \), as in Section 32.4. In particular, every automorphism of \( \Phi \) maps \( E_\mathbb{Q} \) onto itself.

As before, we let \( V \) be a vector space over \( k \) that is a module over \( A \), as a Lie algebra over \( k \), with \( V \) irreducible as a module over \( A \), and standard cyclic with weight \( \mu \in A'_0 \), and suppose that \( \mu(h_\alpha) \) corresponds to a nonnegative integer for each \( \alpha \in \Delta \), with respect to the standard embedding of \( \mathbb{Q} \) into \( k \). If \( \nu \in A'_0 \) is a weight of \( V \) with respect to \( A_0 \), then it follows that \( \nu(h_\alpha) \) corresponds to an integer for every \( \alpha \in \Delta \), as in the previous section. This implies that \( \nu \in E_\mathbb{Q} \), as in Section 32.4.

If \( \alpha \in \Phi \), then we let \( \sigma_\alpha \) be the symmetry on \( E_\mathbb{R} \) with vector \( \alpha \) that maps \( \Phi \) onto itself, and \( \lambda_\alpha \) be the corresponding linear functional on \( E_\mathbb{R} \), as usual. Remember that \( \lambda_\alpha \) maps \( E_\mathbb{Q} \) into \( \mathbb{Q} \), because \( \lambda_\alpha \) takes integer values on \( \Phi \). If \( \nu \in E_\mathbb{Q} \subseteq A'_0 \), then

\[
\lambda_\alpha(\nu) = \nu(h_\alpha),
\]

by (31.1.2), and using the natural embedding of \( \mathbb{Q} \) into \( k \). This implies that

\[
\sigma_\alpha(\nu) = \nu - \lambda_\alpha(\nu) \alpha = \nu - \nu(h_\alpha) \alpha.
\]

If \( \nu \in E_\mathbb{Q} \) and \( \sigma \) is in the Weyl group of \( \Phi \), then we would like to check that \( \nu \) is a weight of \( V \) with respect to \( A_0 \) if and only if \( \nu(\sigma) \) has this property. If \( \sigma = \sigma_\alpha \) for some \( \alpha \in \Delta \), then this follows from the remarks in the previous section, because (32.8.2) is the same as (32.7.4). Otherwise, one can use the fact that the Weyl group of \( \Phi \) is generated by \( \sigma_\alpha \), \( \alpha \in \Delta \), as in Section 19.14. Similarly,

\[
\dim V_\sigma = \dim V_{\sigma(\nu)}
\]

for every \( \sigma \) in the Weyl group of \( \Phi \). More precisely, this is the same as (32.7.6) when \( \sigma = \sigma_\alpha \) for some \( \alpha \in \Delta \).

The weights of \( V \) with respect to \( A_0 \) are elements of \( E_\mathbb{Q} \), as before. The Weyl group of \( \Phi \) maps the set of weights of \( V \) with respect to \( A_0 \) onto itself, as in the preceding paragraph. This and (32.8.3) correspond to step (7) on p114 of [13], and are included in the theorem on p113 of [13]. These statements also correspond to Remark (1) on p62 of [23].

We would like to show that there are only finitely many weights of \( V \) with respect to \( A_0 \). If \( \nu \) is a weight of \( V \), then \( \nu \) can be expressed as \( \mu \) minus a linear
combination of the elements of $\Delta$ with nonnegative integer coefficients, as in Section 31.8. This means that $\nu \leq \mu$, in the notation of Section 30.5.

If $\nu$ is a weight of $V$, then $\nu(h_\alpha)$ corresponds to an integer for every $\alpha \in \Delta$, so that $\nu \in \mathcal{T}$, as in Section 32.4. In this case, there is an element $\sigma$ of the Weyl group of $\Phi$ such that $\sigma(\nu) \in \mathcal{T}^+$, as in Section 30.4. Note that $\sigma(\nu)$ is a weight of $V$ as well, so that $\sigma(\nu) \leq \mu$, as before.

However, there are only finitely many elements $\rho$ of $\mathcal{T}^+$ such that $\rho \leq \mu$, as in Section 30.6. This implies that there are only finitely many elements of the Weyl group of $\Phi$. This corresponds to step (8) on p114 of [13].

Alternatively, if $\nu$ is a weight of $V$, then $\nu$ can be expressed as $\mu$ minus a linear combination of the elements of $\Delta$ with nonnegative integer coefficients, as before. In order to show that there are only finitely many weights of $V$, it suffices to find upper bounds for these coefficients that do not depend on $\nu$, as on p61 of [23].

Of course, $\mu \in E_\mathbb{Q}$, because $\mu(h_\alpha)$ corresponds to an integer for every $\alpha \in \Delta$, by hypothesis. Remember that $\Delta$ is a basis for $E_\mathbb{Q}$ as a vector space over $\mathbb{Q}$, as in Section 31.1. Thus $\nu$ can be expressed in a unique way as a linear combination of elements of $\Delta$ with coefficients in $\mathbb{Q}$, and we would like to find lower bounds for these coefficients that do not depend on $\nu$.

Remember that $-\Delta$ is a base for $\Phi$ too, so that there is an element $\sigma$ of the Weyl group of $\Phi$ such that $\sigma(\Delta) = -\Delta$, as in Section 19.14. We also have that $\sigma(\nu)$ is a weight of $V$, because $\nu$ is a weight of $V$, as before. This means that $\sigma(\nu)$ can be expressed as $\mu$ minus a linear combination of elements of $\Delta$ with nonnegative integer coefficients, as in Section 31.8 again.

It follows that $\nu$ can be expressed as $\sigma^{-1}(\mu)$ minus a linear combination of the elements of $\sigma^{-1}(\Delta)$ with nonnegative integer coefficients. Equivalently, $\nu$ can be expressed as $\sigma^{-1}(\mu)$ plus a linear combination of elements of $\Delta$ with nonnegative integer coefficients. In particular, the coefficients of $\nu$ with respect to $\Delta$ are greater than or equal to the coefficients of $\sigma^{-1}(\mu)$.

This implies that the coefficients of $\nu$ with respect to $\Delta$ have lower bounds that do not depend on $\nu$, because there are only finitely many elements of the Weyl group. It follows that there are only finitely many weights of $V$, as before.

Remember that $V$ is spanned by its weight spaces, each of which has finite dimension, as in Section 31.8. This means that $V$ has finite dimension, because there are only finitely many weights of $V$. This corresponds to step (9) on p114 of [13], and to the last step in the proof on p61 of [23].

### 32.9 Sums of weights

Let us continue to use the same notation and hypotheses as in Section 31.1. Let $V^1, V^2$ be vector spaces over $k$ that are modules over $A$, as a Lie algebra over $k$. Consider the tensor product $V^1 \otimes V^2$, initially as a vector space over $k$. As in Section 7.12, $V^1 \otimes V^2$ may be considered as a module over $A$, where

\[
(32.9.1) \quad a \cdot (z_1 \otimes z_2) = (a \cdot z_1) \otimes z_2 + z_1 \otimes (a \cdot z_2)
\]
32.9. SUMS OF WEIGHTS

for every \( a \in A \), \( z_1 \in V^1 \), and \( z_2 \in V^2 \).

Let \( \nu_1, \nu_2 \in A'_0 \) be given, and remember that \( V^1_{\nu_1}, V^2_{\nu_2} \) consist of elements of \( V^1, V^2 \) with weight \( \nu_1, \nu_2 \), respectively, with respect to \( A_0 \), as in Section 31.5. If \( w \in A_0, z_1 \in V^1_{\nu_1}, \) and \( z_2 \in V^2_{\nu_2} \), then

\[
(32.9.2) \quad w \cdot (z_1 \otimes z_2) = (w \cdot z_1) \otimes z_2 + z_1 \otimes (w \cdot z_2) = (\nu_1(w) z_1) \otimes z_2 + z_1 \otimes (\nu_2(w) z_2) = (\nu_1(w) + \nu_2(w)) (z_1 \otimes z_2).
\]

This means that \( z_1 \otimes z_2 \) is in the set \((V^1 \otimes V^2)_{\nu_1+\nu_2}\) of elements of \( V^1 \otimes V^2 \) with weight \( \nu_1 + \nu_2 \) with respect to \( A_0 \). The tensor product \( V^1_{\nu_1} \otimes V^2_{\nu_2} \) may be considered as a linear subspace of \( V^1 \otimes V^2 \), and we get that

\[
(32.9.3) \quad V^1_{\nu_1} \otimes V^2_{\nu_2} \subseteq (V^1 \otimes V^2)_{\nu_1+\nu_2}.
\]

Let \( \mu_1, \mu_2 \in A'_0 \) be given, and suppose that

\[
(32.9.4) \quad v_1 \in V^1, v_2 \in V^2 \text{ are primitive or maximal vectors of weight } \mu_1, \mu_2, \text{ respectively,}
\]

as in Section 31.6. Under these conditions, it is easy to see that

\[
(32.9.5) \quad v_1 \otimes v_2 \text{ is a primitive or maximal vector in } V^1 \otimes V^2 \text{ of weight } \mu_1 + \mu_2.
\]

If \( UA \) is a universal enveloping algebra of \( A \), then

\[
(32.9.6) \quad V = (UA) \cdot (v_1 \otimes v_2)
\]

is the submodule of \( V^1 \otimes V^2 \) as a module over \( A \), generated by \( v_1 \otimes v_2 \), as before. Thus

\[
(32.9.7) \quad V \text{ is standard cyclic of weight } \mu_1 + \mu_2,
\]

as a module over \( A \).

If \( V^1, V^2 \) have finite dimension as vector spaces over \( k \), then \( V^1 \otimes V^2 \) has finite dimension as well, which implies that \( V \) has finite dimension too. This shows that

\[
(32.9.8) \quad \text{if there are standard cyclic modules over } A \text{ with weights } \mu_1, \mu_2 \text{ that are finite dimensional as vector spaces over } k,
\]

then there is a standard cyclic module over \( A \) with weight \( \mu_1 + \mu_2 \) that is finite dimensional as a vector space over \( k \),

as in Proposition 4.4 on p61 of [24]. Of course, this also follows from the characterization of weights of finite-dimensional standard cyclic modules stated in Section 32.3. This more direct argument for sums is used in [24] to obtain the sufficiency of the condition in Section 32.3 when \( A = sl_n(k) \).
32.10 Weights and tensor products

Let us continue with the same notation and hypotheses as in the previous section. If $V$ is a vector space over $k$ that is a module over $A$, then put

$$
\Pi(V) = \{ \nu \in A_0' : V_\nu \neq \{0\} \},
$$

which is the set of weights of $V$ with respect to $A_0$. Observe that

$$
\{ \nu_1 + \nu_2 : \nu_1 \in \Pi(V^1), \nu_2 \in \Pi(V^2) \} \subseteq \Pi(V^1 \otimes V^2),
$$

by (32.9.3).

Suppose that $V^1, V^2$ correspond to the direct sums of their weight spaces, as vector spaces over $k$:

This implies that

$$
V^1 \otimes V^2 \text{ corresponds to the direct sum of } V^1_{\nu_1} \otimes V^2_{\nu_2},
$$

$\nu_1 \in \Pi(V^1), \nu_2 \in \Pi(V^2)$, as a vector space over $k$,

as in Section 7.14.

Under these conditions, we get that

$$
\Pi(V^1 \otimes V^2) = \{ \nu_1 + \nu_2 : \nu_1 \in \Pi(V^1), \nu_2 \in \Pi(V^2) \},
$$

using (32.9.3). More precisely, if $\nu \in A_0'$, then

$$
(V^1 \otimes V^2)_\nu \text{ corresponds to the direct sum of } V^1_{\nu_1} \otimes V^2_{\nu_2},
$$

where $\nu_1 \in \Pi(V^1), \nu_2 \in \Pi(V^2)$, and $\nu_1 + \nu_2 = \nu$,

as a vector space over $k$.

Note that

$$
V^1 \otimes V^2 \text{ corresponds to the direct sum of its weight spaces, as a vector space over } k,
$$

in this case.

If $V^1, V^2$ have finite dimension as vector spaces over $k$, then $\Pi(V^1), \Pi(V^2)$ have only finitely many elements. This implies that $\Pi(V^1 \otimes V^2)$ has only finitely many elements as well, by (32.10.5). If $\nu \in \Pi(V^1 \otimes V^2)$, then we get that

$$
\dim (V^1 \otimes V^2)_\nu = \sum_{\nu_1 + \nu_2 = \nu} (\dim V^1_{\nu_1}) (\dim V^2_{\nu_2}),
$$

where more precisely the sum is taken over all $\nu_1 \in \Pi(V^1)$ and $\nu_2 \in \Pi(V^2)$ with $\nu_1 + \nu_2 = \nu$. 

Suppose now that\( V^1, V^2 \) are standard cyclic of weights \( \mu_1, \mu_2 \in A'_0 \), respectively, as modules over \( A \).

In particular, this implies that (32.10.3) holds, as in Section 31.8. Remember that the elements of \( \Pi(V^1), \Pi(V^2) \) can be expressed as \( \mu_1, \mu_2 \) minus linear combinations of elements of \( \Delta \) with nonnegative integer coefficients, respectively, as in Section 31.8. It follows that

\[
\text{(32.10.10)} \quad \text{the elements of } \Pi(V^1 \otimes V^2) \text{ can be expressed as} \quad \mu_1 + \mu_2 \text{ minus linear combinations of elements of } \Delta \\
\text{with nonnegative integer coefficients.}
\]

If \( \nu_1 \in \Pi(V^1), \nu_2 \in \Pi(V^2) \), then we get that

\[
\text{(32.10.11)} \quad \nu_1 + \nu_2 = \mu_1 + \mu_2 \quad \text{only when} \quad \nu_1 = \mu_1, \nu_2 = \mu_2.
\]

This implies that

\[
\text{(32.10.12)} \quad (V^1 \otimes V^2)_{\mu_1 + \mu_2} = V^1_{\mu_1} \otimes V^2_{\mu_2},
\]

as in (32.10.6). Remember that \( V^1_{\mu_1}, V^2_{\mu_2} \) have dimension one as vector spaces over \( k \), as in Section 31.8. This means that

\[
\text{(32.10.13)} \quad \dim (V^1 \otimes V^2)_{\mu_1 + \mu_2} = 1,
\]

by (32.10.12).

If \( \nu \in \Pi(V^1 \otimes V^2) \), then one can check that

\[
\text{(32.10.14)} \quad \nu = \nu_1 + \nu_2 \text{ for only finitely many } \nu_1 \in \Pi(V^1), \nu_2 \in \Pi(V^2),
\]

because \( \nu_1, \nu_2 \) can be expressed as \( \mu_1, \mu_2 \) minus linear combinations of elements of \( \Delta \) with nonnegative integer coefficients, as before. Remember that \( V^1_{\nu_1}, V^2_{\nu_2} \) have finite dimensions as vector spaces over \( k \) for every \( \nu_1, \nu_2 \in A'_0 \), as in Section 31.8. Thus

\[
\text{(32.10.15)} \quad (V^1 \otimes V^2)_{\nu} \text{ has finite dimension}
\]
as a vector space over \( k \), because it corresponds to the direct sum of finitely many finite-dimensional subspaces, by (32.10.6). More precisely, the dimension of \( (V^1 \otimes V^2)_{\nu} \) can be expressed as in (32.10.8) again.

This corresponds to Exercise 7 on p117 of [13], and to part of part (b) of Proposition 5 on p63 of [23].

### 32.11 Fundamental weights and modules

Let us return to some connections with abstract weights, as in Section 32.4. Remember that the collection \( \Phi' \) of \( \lambda_\alpha, \alpha \in \Phi \), is a root system in the dual \( E'_R \)
of $E_R$, and that the set $\Delta'$ of $\lambda_\alpha$, $\alpha \in \Delta$, is a base for $\Phi'$, because $\Phi$ is reduced as a root system in $E_R$. In particular, $\Delta'$ is a basis for $E'_R$, as a vector space over $R$, and if $\alpha \in \Delta$, then there is a unique element $\mu_\alpha$ of $E_R$ such that

$$\lambda_\beta(\mu_\alpha) = \begin{cases} 1 & \text{when } \beta = \alpha \\ 0 & \text{for every } \beta \in \Delta \text{ with } \beta \neq \alpha, \end{cases}$$

as in Section 30.1. The $\mu_\alpha$’s, $\alpha \in \Delta$, form a basis for $E_R$ dual to the $\lambda_\beta$’s, $\beta \in \Delta$, as before.

Remember that $\Upsilon$ consists of $z \in E_R$ such that $\lambda_\beta(z) \in \mathbb{Z}$ for every $\beta \in \Phi$, or equivalently for every $\beta \in \Delta$, because $\Delta'$ is a base for $\Phi'$. Thus $\mu_\alpha \in \Upsilon$ for every $\alpha \in \Delta$, as in Section 30.1. We also have that $\Upsilon \subset E_Q$, as in Section 32.4, so that $\mu_\alpha \in E_Q \subset A'_0$ for every $\alpha \in \Delta$.

Similarly, $\Upsilon^+$ consists of $z \in \Upsilon$ such that $\lambda_\beta(z) \geq 0$ for every $\beta \in \Delta$, or equivalently for every $\beta \in \Phi$. Clearly $\mu_\alpha \in \Upsilon^+$ for every $\alpha \in \Delta$, as in Section 30.3. Remember that the $\mu_\alpha$’s, $\alpha \in \Delta$, are called the fundamental (dominant) weights with respect to $\Delta$, as in Section 30.1. It is easy to see that $\Upsilon^+$ consists exactly of linear combinations of the $\mu_\alpha$’s, $\alpha \in \Delta$, whose coefficients are nonnegative integers.

As in Section 32.4, $\Upsilon$ consists exactly of $\nu \in A'_0$ such that $\nu(\lambda_\beta)$ corresponds to an integer for every $\beta \in \Delta$, or equivalently for every $\beta \in \Phi$. Similarly, $\Upsilon^+$ consists exactly of $\mu \in A'_0$ such that $\mu(\lambda_\beta)$ corresponds to a nonnegative integer for every $\beta \in \Delta$, or equivalently $\beta \in \Phi$. This is the same as the set of $\mu \in A'_0$ for which there is a finite-dimensional standard cyclic module over $A$ with weight $\mu$, as in Section 32.3. The modules associated to $\mu_\alpha$, $\alpha \in \Delta$, in this way are called the fundamental modules over $A$, or the fundamental representations of $A$, as in Remark (2) on p62 of [23].

If $\mu \in \Upsilon^+$, then $\mu$ can be expressed as a sum of $\mu_\alpha$’s, $\alpha \in \Delta$, with suitable repetitions. One can use this to get a finite-dimensional standard cyclic module over $A$ with weight $\mu$ as a submodule of a tensor product of fundamental modules over $A$, with suitable repetitions, as in Exercise 8 on p117 of [13]. This is analogous to the argument discussed in Section 32.9. This type of argument is used on p61 of [24], to get the existence of finite-dimensional standard cyclic modules over $A = sl_n(k)$ with any weight $\mu \in \Upsilon^+$.

### 32.12 Simple transitivity

Let us continue with the same notation and hypotheses as in Section 31.1. In particular, $\Phi$ is a reduced root system in the vector space $E_R$ over $R$ of positive finite dimension, and $\Delta$ is a base for $\Phi$. Remember that any reduced root system can be as in Section 31.1, because of Serre’s theorem, as in Section 27.8.

If $\sigma$ is an element of the Weyl group of $\Phi$ such that

$$\sigma(\Delta) = \Delta,$$

then $\sigma$ is the identity mapping on $E_R$. This was discussed in Section 20.1, and it can also be obtained using results about representations of the Lie algebra $A$,
as in Proposition 4 on p62 of [23]. This corresponds to Exercise 1 on p116 of [13].

If \( \beta \in \Phi \), then
\[
\lambda_{\sigma(\beta)} = \lambda_{\beta} \circ \sigma^{-1},
\]
(32.12.2)

because \( \sigma \) is an automorphism of \( \Phi \), as in Section 19.4. Let \( \mu_{\alpha}, \alpha \in \Delta \), be the fundamental weights with respect to \( \Delta \), as in the previous section. If \( \alpha, \beta \in \Delta \), then
\[
\lambda_{\sigma(\beta)}(\sigma(\mu_{\alpha})) = \lambda_{\beta}(\mu_{\alpha}) = \begin{cases} 1 & \text{when } \beta = \alpha \\ 0 & \text{when } \beta \neq \alpha, \end{cases}
\]
(32.12.3)

by (32.12.2) and (32.11.1). This implies that
\[
\sigma(\mu_{\alpha}) = \mu_{\sigma(\alpha)}
\]
(32.12.4)

for every \( \alpha \in \Delta \), because of (32.12.1).

Let \( \alpha \in \Delta \) be given, and let \( V^\alpha \) be a finite-dimensional standard cyclic module over \( A \) with weight \( \mu_{\alpha} \). In particular, \( \mu_{\alpha} \) is a weight of \( V^\alpha \), which implies that (32.12.4) is a weight of \( V^\alpha \) too, as in Section 32.8. It follows that
\[
\mu_{\alpha} - \sigma(\mu_{\alpha}) = \mu_{\alpha} - \mu_{\sigma(\alpha)}
\]
(32.12.5)

can be expressed as a linear combination of the elements of \( \Delta \), with coefficients that are nonnegative integers, as in Section 31.8.

Of course, the sum of (32.12.5) over \( \alpha \in \Delta \) is equal to 0, by (32.12.1). This implies that (32.12.5) is equal to 0 for every \( \alpha \in \Delta \), because \( \Delta \) is a basis for \( E_R \), and the coefficients of (32.12.5) with respect to \( \Delta \) are all nonnegative. It follows that \( \sigma \) is the identity mapping on \( E_R \), as desired.

### 32.13 Weights and saturated sets

We continue with the same notation and hypotheses as in Section 31.1. Let \( V \) be a finite-dimensional vector space over \( k \), and suppose that \( V \) is a standard cyclic module over \( A \) of weight \( \mu \in A_0' \). If \( \nu \in A_0' \), then let \( V_\nu \) be the linear subspace of \( V \) consisting of vectors of weight \( \nu \) with respect to \( A_0 \), as in Section 31.5. Put
\[
\Pi = \{ \nu \in A_0' : V_\nu \neq \{0\} \}.
\]
(32.13.1)

Note that \( \Pi \) has only finitely many elements, because \( V \) has finite dimension, by hypothesis.

If \( \nu \in \Pi \) and \( \alpha \in \Phi \), then \( \nu(h_\alpha) \) corresponds to an integer under the standard embedding of \( Q \) into \( k \), as in Sections 32.1 and 32.2. This means that
\[
\Pi \subseteq \Upsilon,
\]
(32.13.2)

where \( \Upsilon \) is as in Section 32.4. We would like to show that
\[
\Pi \text{ is saturated as a subset of } \Upsilon,
\]
(32.13.3)
in the sense of Section 30.7. This corresponds to the first part of the proposition on page 114 of [13].

Let \( \nu \in \Pi \) and \( \alpha \in \Phi \) be given, and remember that \( \nu \in E_{\mathbb{Q}} \) and \( \nu(h_{\alpha}) = \lambda_{\alpha}(\nu) \), as in Section 32.4. Let \( x_{\alpha} \) be a nonzero element of \( A_{\alpha} \), and let \( y_{\alpha} \) be an element of \( A_{-\alpha} \) such that \( [x_{\alpha}, y_{\alpha}]_A = h_{\alpha} \), as usual. The linear span of \( x_{\alpha}, y_{\alpha}, \) and \( h_{\alpha} \) in \( A \) is a Lie subalgebra of \( A \) that is isomorphic to \( \mathfrak{sl}_2(\mathbb{k}) \), as a Lie algebra over \( \mathbb{k} \).

Let \( Z \) be the linear subspace of \( V \) spanned by the subspaces of the form

\[
V_{\nu + j \cdot \alpha}, \quad j \in \mathbb{Z}.
\]

It is easy to see that \( Z \) is a submodule of \( V \), as a module over the linear span of \( x_{\alpha}, y_{\alpha}, \) and \( h_{\alpha} \) in \( A \), using the remarks in Section 31.5.

Suppose that \( j_1, j_2 \in \mathbb{Z} \) satisfy \( j_1 \leq j_2 \) and

\[
\nu + j_1 \cdot \alpha, \quad \nu + j_2 \cdot \alpha \in \Pi.
\]

If \( j \in \mathbb{Z} \) and \( j_1 \leq j \leq j_2 \), then we would like to verify that

\[
\nu + j \cdot \alpha \in \Pi.
\]

Of course, \( Z \) has finite dimension as a vector space over \( \mathbb{k} \), because \( V \) has finite dimension, by hypothesis. Thus we can use Weyl’s theorem to get that \( Z \) corresponds to the direct sum of irreducible modules over the linear span of \( x_{\alpha}, y_{\alpha}, \) and \( h_{\alpha} \) in \( A \). Each of these irreducible modules is isomorphic to \( W(m) \) in Section 15.4 for some nonnegative \( m \), as a module over \( \mathfrak{sl}_2(\mathbb{k}) \), as in Sections 15.5 and 15.9. One can use this to get (32.13.6), because of the analogous property of \( W(m) \).

Remember that elements of the Weyl group of \( \Phi \) map \( \Pi \) onto itself, as in Section 32.8. In particular, \( \sigma_{\alpha}(\nu) \in \Pi \), where \( \sigma_{\alpha} \) is the symmetry on \( E_{\mathbb{R}} \) with vector \( \alpha \) that maps \( \Phi \) onto itself, as usual. Equivalently,

\[
\nu - \lambda_{\alpha}(\nu) \alpha \in \Pi,
\]

where \( \lambda_{\alpha}(\nu) = \nu(h_{\alpha}) \) corresponds to an integer under the natural embedding of \( \mathbb{Q} \) into \( \mathbb{k} \), as before. If \( j \in \mathbb{Z} \) is between 0 and \( \lambda_{\alpha}(\nu) \), then (32.13.6) holds, because \( \nu \in \Pi \), by hypothesis. This implies (32.13.3), as desired.

Let \( r, q \) be the largest integers such that

\[
\nu - r \cdot \alpha, \quad \nu + q \cdot \alpha \in \Pi,
\]

so that \( q, r \geq 0 \), because \( \nu \in \Pi \). It follows that (32.13.6) holds for some \( j \in \mathbb{Z} \) if and only if \( -r \leq j \leq q \). The sequence of elements of \( \Pi \) of the form

\[
\nu + j \cdot \alpha,
\]

with \( j \in \mathbb{Z} \) and \( -r \leq j \leq q \), is called the \( \alpha \)-string through \( \nu \), as on page 114 of [13]. It is easy to see that \( \sigma_{\alpha} \) maps this sequence onto itself, but in the opposite order. This implies that

\[
r - q = \lambda_{\alpha}(\nu),
\]
as on p114 of [13].

Of course, \( \mu \in \Pi \), and \( \mu(h_\alpha) \) corresponds to a nonnegative integer for every \( \alpha \in \Phi^+ \), as in Section 32.3. This means that \( \mu \in \Upsilon^+ \), as in Section 32.4. In fact,

\[
(32.13.11) \quad \Pi \text{ has highest weight } \mu,
\]

in the sense of Section 30.7. Indeed, if \( \nu \in \Pi \), then \( \nu \) can be expressed as \( \mu \) minus a linear combination of elements of \( \Delta \) with nonnegative integer coefficients, as in Section 31.8. Thus \( \nu \preceq \mu \), in the notation of Section 30.5.

If \( \rho \in \Upsilon^+ \) and \( \rho \preceq \mu \), then

\[
(32.13.12) \quad \rho \in \Pi,
\]

as in Section 30.8. This implies that

\[
(32.13.13) \quad \sigma(\rho) \in \Pi
\]

for every element \( \sigma \) of the Weyl group of \( \Phi \), as before.

If \( \nu \in \Pi \) and \( \sigma \) is an element of the Weyl group of \( \Phi \), then \( \sigma(\nu) \in \Pi \), so that

\[
(32.13.14) \quad \sigma(\nu) \preceq \mu,
\]

as before. Conversely, if \( \nu \in \Upsilon \) and (32.13.14) holds for every element \( \sigma \) of the Weyl group of \( \Phi \), then \( \nu \in \Pi \). Indeed, if \( \nu \in \Upsilon \), then there is an element \( \sigma \) of the Weyl group of \( \Phi \) such that \( \sigma(\nu) \in \Upsilon^+ \), as in Section 30.4. This implies that \( \sigma(\nu) \in \Pi \), and thus \( \nu \in \Pi \), as in the preceding paragraph. This corresponds to the second part of the proposition on p114 of [13].
**Bibliography**


Index

\(A((T)),\) 163
\(A[[T]],\) 81, 163
\(A[[T_1,\ldots,T_n]]\), 111
\(A[T],\) 81
\(A[T_1,\ldots,T_n],\) 111
\(A^{op},\) 138
\(A_X(k) = A_X,\) 555
\(A_X^a(k) = A_X^a,\) 555
absolute convergence, 19
absolute value functions, 3
ultrametric, 4
absolute values, 181
absolutely simple Lie algebras, 330
absorbing sets, 22
abstract Jordan decompositions, 289
abstract weights, 630
ad-diagonalizable elements, 288
ad-nilpotent elements, 193
ad-semisimple elements, 288
adjoint mappings, 18, 46, 48, 65, 67, 234, 251
algebra automorphisms, 28, 34, 35
algebra homomorphisms, 28, 34, 35
algebra isomorphisms, 28, 34, 35
algebras, 27
associated to magmas, 555
\(\alpha\)-string through \(\beta,\) 410
\(\alpha\)-string through \(\nu,\) 690
anti-self-adjointness, 34, 73, 321
linear mappings, 47, 250
antisymmetric mappings,
  bilinear, 27, 43
  linear, 43
  multilinear, 145
archimedean absolute value functions, 4
archimedean property, 182

Assoc\(_X(k) = \text{Assoc}_X,\) 559
Assoc\(_X^a(k) = \text{Assoc}_X^a,\) 561
associative algebras, 28
associative bilinear mappings, 136
balanced sets, 22
balanced subgroups, 24
Banach algebras, 58
Banach spaces, 12
bases for free modules, 539
bases of root systems, 392
bilinear actions, 123
bilinear forms, 43
bilinear mappings, 20, 27
\(\mathcal{B}\mathcal{L}(V),\) 52
\(\mathcal{B}\mathcal{L}(V,W),\) 13
\(\mathcal{B}\mathcal{L}(V_1,\ldots,V_n;W),\) 146
Borel subalgebras, 465, 473
bounded bilinear mappings, 20
bounded linear mappings, 13
bounded multilinear mappings, 146
C, 3
c\((X,W),\) 159
c\(_{00}(X,W),\) 159
\(C_A(E),\) 149
Cartan matrix, 406
Cartan subalgebras, 368
Casimir elements, 274, 276, 277
Cauchy products, 79
Cauchy–Schwarz inequality, 17
center, 149, 150
centralizer, 149
Chevalley algebras, 624
Chevalley bases, 615
closed balls, 3
commutative algebras, 28
INDEX

fundamental group (root systems), 633
fundamental modules, 688
fundamental representations, 688
fundamental root systems, 392
fundamental weights, 631

general linear algebra, 39, 40
general linear group, 51
generators and relations, 567
$GL(V)$, 51
$gl(V)$, 40
$GL_n(A)$, 51
$gl_n(A)$, 39
graded algebra associated to $UA$, 538

Hermitian forms, 47, 250
highest weights, 640, 664
Hilbert spaces, 17
$\text{Hom}(A, B)$, 27
$\text{Hom}_k(A, B)$, 27

homomorphisms
between algebras, 28, 34, 35
between Lie algebras, 30
between modules, 26, 131, 132
between representations, 131, 132

ideals, 41, 42
identity matrix, 38
indecomposability, 661
inner automorphism, 300
inner derivations, 285
inner products, 16, 237, 251
Int $A$, 299
integral weights, 677
intertwining mappings, 125, 131, 132
invariant bilinear mappings, 64
invariant elements, 150
invariant sesquilinear forms, 67
inverse roots, 387
inverse systems, 388
invertible elements, 50
invertible linear mappings, 60
involutions, 34, 35, 322
irreducible representations, 141
irreducible root systems, 409
isometric linear mappings, 60

diagonal automorphisms, 524, 597, 613
diagonal matrices, 549
diagonalizable linear mappings, 213
direct products, 77
direct sums, 77
discrete absolute value functions, 7
diagonal automorphisms, 524, 597, 613
diagonal matrices, 209
diagonalizable linear mappings, 213

fundamental group (root systems), 633
fundamental modules, 688
fundamental representations, 688
fundamental root systems, 392
fundamental weights, 631

Hermitian forms, 47, 250
highest weights, 640, 664
Hilbert spaces, 17
$\text{Hom}(A, B)$, 27
$\text{Hom}_k(A, B)$, 27

homomorphisms
between algebras, 28, 34, 35
between Lie algebras, 30
between modules, 26, 131, 132
between representations, 131, 132

ideals, 41, 42
identity matrix, 38
indecomposability, 661
inner automorphism, 300
inner derivations, 285
inner products, 16, 237, 251
Int $A$, 299
integral weights, 677
intertwining mappings, 125, 131, 132
invariant bilinear mappings, 64
invariant elements, 150
invariant sesquilinear forms, 67
inverse roots, 387
inverse systems, 388
invertible elements, 50
invertible linear mappings, 60
involutions, 34, 35, 322
irreducible representations, 141
irreducible root systems, 409
isometric linear mappings, 60

diagonal automorphisms, 524, 597, 613
diagonal matrices, 209
diagonalizable linear mappings, 213
direct products, 77
direct sums, 77
discrete absolute value functions, 7
diagonal automorphisms, 524, 597, 613
diagonal matrices, 209
diagonalizable linear mappings, 213
direct products, 77
direct sums, 77
discrete absolute value functions, 7
diagonal automorphisms, 524, 597, 613
diagonal matrices, 209
diagonalizable linear mappings, 213
direct products, 77
direct sums, 77
discrete absolute value functions, 7

fundamental group (root systems), 633
fundamental modules, 688
fundamental representations, 688
fundamental root systems, 392
fundamental weights, 631

Hermitian forms, 47, 250
highest weights, 640, 664
Hilbert spaces, 17
$\text{Hom}(A, B)$, 27
$\text{Hom}_k(A, B)$, 27

homomorphisms
between algebras, 28, 34, 35
between Lie algebras, 30
between modules, 26, 131, 132
between representations, 131, 132

ideals, 41, 42
identity matrix, 38
indecomposability, 661
inner automorphism, 300
inner derivations, 285
inner products, 16, 237, 251
Int $A$, 299
integral weights, 677
intertwining mappings, 125, 131, 132
invariant bilinear mappings, 64
invariant elements, 150
invariant sesquilinear forms, 67
inverse roots, 387
inverse systems, 388
invertible elements, 50
invertible linear mappings, 60
involutions, 34, 35, 322
irreducible representations, 141
irreducible root systems, 409
isometric linear mappings, 60

diagonal automorphisms, 524, 597, 613
diagonal matrices, 209
diagonalizable linear mappings, 213
direct products, 77
direct sums, 77
discrete absolute value functions, 7
diagonal automorphisms, 524, 597, 613
diagonal matrices, 209
diagonalizable linear mappings, 213
direct products, 77
direct sums, 77
discrete absolute value functions, 7

fundamental group (root systems), 633
fundamental modules, 688
fundamental representations, 688
fundamental root systems, 392
fundamental weights, 631

Hermitian forms, 47, 250
highest weights, 640, 664
Hilbert spaces, 17
$\text{Hom}(A, B)$, 27
$\text{Hom}_k(A, B)$, 27

homomorphisms
between algebras, 28, 34, 35
between Lie algebras, 30
between modules, 26, 131, 132
between representations, 131, 132

ideals, 41, 42
identity matrix, 38
indecomposability, 661
inner automorphism, 300
inner derivations, 285
inner products, 16, 237, 251
Int $A$, 299
integral weights, 677
intertwining mappings, 125, 131, 132
invariant bilinear mappings, 64
invariant elements, 150
invariant sesquilinear forms, 67
inverse roots, 387
inverse systems, 388
invertible elements, 50
invertible linear mappings, 60
involutions, 34, 35, 322
irreducible representations, 141
irreducible root systems, 409
isometric linear mappings, 60

diagonal automorphisms, 524, 597, 613
diagonal matrices, 209
diagonalizable linear mappings, 213
direct products, 77
direct sums, 77
discrete absolute value functions, 7
diagonal automorphisms, 524, 597, 613
diagonal matrices, 209
diagonalizable linear mappings, 213
direct products, 77
direct sums, 77
discrete absolute value functions, 7

fundamental group (root systems), 633
fundamental modules, 688
fundamental representations, 688
fundamental root systems, 392
fundamental weights, 631

Hermitian forms, 47, 250
highest weights, 640, 664
Hilbert spaces, 17
$\text{Hom}(A, B)$, 27
$\text{Hom}_k(A, B)$, 27

homomorphisms
between algebras, 28, 34, 35
between Lie algebras, 30
between modules, 26, 131, 132
between representations, 131, 132

ideals, 41, 42
identity matrix, 38
indecomposability, 661
inner automorphism, 300
inner derivations, 285
inner products, 16, 237, 251
Int $A$, 299
integral weights, 677
intertwining mappings, 125, 131, 132
invariant bilinear mappings, 64
invariant elements, 150
invariant sesquilinear forms, 67
inverse roots, 387
inverse systems, 388
invertible elements, 50
invertible linear mappings, 60
involutions, 34, 35, 322
irreducible representations, 141
irreducible root systems, 409
isometric linear mappings, 60
isomorphic ordered fields, 183
isomorphic representations, 131, 132
isomorphic root systems, 382
Jacobi identity, 29
kernels of module homomorphisms, 41
killing form, 154
$L(V)$, 40
$L(V, W)$, 40
$L(V_1, \ldots, V_n; W)$, 144
$L_X(k) = L_X$, 556
$L^X(k) = L^X$, 561
lengths in the Weyl group, 446
lengths of $m$-tuples, 540
Lie algebra homomorphisms, 30
Lie algebras, 30
Lie bracket, 30
Lie subalgebras, 30
Lie’s theorem, 202
linear mappings, 27
locally nilpotent linear mappings, 584
long roots, 414
lower central series, 190
$M_n(A)$, 37
$M_X$, 554
magma homomorphisms, 554
magmas, 554
maximal roots, 415
maximal vector of weight $\mu$, 657
maximal vector of weight $\lambda$, 306
metrics, 2
module automorphisms, 64
module homomorphisms, 26, 131, 132
module isomorphisms, 27, 131, 132
modules,
  over associative algebras, 128
  over $k$, 26
over Lie algebras, 129
multi-indices, 110, 526
multilinear mappings, 144
multiplicative identity elements, 28
multiplicative seminorms, 52
multiplicities, 655
$N(A)$, 501
nilpotent elements of associative algebras, 193
nilpotent Lie algebras, 190
non-archimedean absolute value functions, 4
nondegenerate bilinear forms, 45
nondegenerate sesquilinear forms, 48, 250
nonnegative self-adjoint operators, 266
normalizer, 194
norms, 11
open balls, 2
open sets, 3
opposite algebra, 138
ordered fields, 180
ordered rings, 180
orthogonal transformations, 62
orthogonal vectors, 237, 251
Ostrowski’s theorems, 6, 7
$PUA$, 551
P. Hall families, 565
$p$-adic absolute value, 4
$p$-adic integers, 9
$p$-adic metric, 4
$p$-adic numbers, 6
poles of bounded order, 164, 166, 169, 170, 172
positive formal series, 182
positive roots, 394
positive self-adjoint operators, 266
primitive elements of $UA$, 549
primitive of weight $\lambda$, 306
primitive of weight $\mu$, 657
products of submodules, 186
$Q$, 4
$Q_p$, 6
$R$, 3
$R_+$, 7
radicals of bilinear mappings, 156
radicals of Lie algebras, 189
ranks of Lie algebras, 520
ranks of root systems, 380
real parts, 258
real-bilinear mappings, 318
real-linear subspace, 317
real-linear mappings, 16, 317
reduced expressions, 446
reduced root systems, 380
reductive Lie algebras, 230
reductions, 378
regular linear functionals, 401
replicas, 220
representations,
of associative algebras, 127
of Lie algebras, 129
residue fields, 10
root lattices, 630
root strings, 410
root systems, 379
type $A_2$, 426, 429
type $A_n$, 418, 426
type $B_n$, 420
type $BC_n$, 421
type $C_n$, 420
type $D_n$, 423
type $E_6$, 443
type $E_7$, 441
type $E_8$, 439
type $F_4$, 434
type $G_2$, 430, 431
roots, 380
$S_n(A)$, 209
$s_n(A)$, 209
saturated sets of weights, 640
scalar matrices, 209
Schur’s lemma, 141
self-adjoint linear mappings, 47, 250
self-adjoint sets, 240, 254
self-adjointness, 34, 73, 321
self-normalizing Lie subalgebras, 195
semi-ultrametrics, 2
semi-ultranorms, 11
semimetrics, 2
seminorms, 11
semisimple Lie algebras, 189
semisimple linear mappings, 288
semisimple modules, 272
Serre’s theorem, 580
sesquilinear forms, 47, 250
short roots, 414
simple Lie algebras, 225
simple modules, 141
simple root systems, 392
simple roots, 392
$SL(V)$, 51
$sl(V)$, 41
$SL_n(A)$, 51
$sl_n(A)$, 39
solvable Lie algebras, 187
special linear algebra, 39, 41
special linear group, 51
standard Borel subalgebras, 473
standard cyclic modules, 659
strong ad-nilpotence, 500
strongly dominant weights, 633
subalgebras, 28, 30, 474
submodules, 26, 130, 131
submultiplicative seminorms, 52
subrepresentations, 130
sum of root systems, 409
supports of functions, 159
symmetric algebras, 536
symmetric mappings,
  bilinear, 27, 43
  linear, 43
  multilinear, 145
symmetric tensors, 162
symmetries, 377
$TV$, 529
$T^nV$, 162
$T_{n,r}(A)$, 199
tensor algebra, 529
tensor powers, 162
tensor products, 158, 666
termwise eventual agreement, 82, 164
termwise eventually constant, 82, 164
toral subalgebras, 335
totally real linear subspaces, 317
trace, 39, 41, 153
transpose matrices, 38
trivial absolute value function, 3
trivial ultranorm, 12

$\mathcal{U}A$, 530
$\mathcal{U}(\mathcal{F})$, 197
ultrametric absolute values, 4
ultrametrics, 2
ultranorms, 11
unitary transformations, 62
universal enveloping algebras, 530
vanishing to some order, 82, 120

weight $\alpha$, 336
weight $\lambda$, 306
weight $\mu$, 655
weight spaces, 306, 655
weights, 306, 655
Weyl chamber, 402
Weyl group, 381
Weyl relations, 485
Weyl’s theorem, 272, 279

$\mathbb{Z}$, 9
$\mathbb{Z}(A)$, 150
$\mathbb{Z}_+$, 4
$\mathbb{Z}_p$, 9