Some topics in analysis related to topological groups and Lie algebras

Stephen Semmes
Rice University
Preface
## Contents

1 Absolute values and norms .......................... 1
   1.1 Metrics and ultrametrics ....................... 1
   1.2 Absolute value functions ...................... 2
   1.3 Equivalent absolute value functions .......... 4
   1.4 Completions ................................ 5
   1.5 Discreteness ................................ 6
   1.6 $p$-Adic integers ............................ 7
   1.7 Residue fields ................................ 9
   1.8 Norms and ultranorms ......................... 10
   1.9 Bounded linear mappings ...................... 12
   1.10 Some norms on $k^n$ ......................... 13
   1.11 Inner products .............................. 15
   1.12 Infinite series ............................. 17
   1.13 Bounded bilinear mappings .................. 19
   1.14 Minkowski functionals ....................... 21
   1.15 Balanced subgroups ......................... 23

2 Some basic notions related to Lie algebras .... 25
   2.1 Modules and homomorphisms .................. 25
   2.2 Algebras ................................ 26
   2.3 Lie algebras ................................ 28
   2.4 The adjoint representation .................. 30
   2.5 Derivations ................................ 31
   2.6 Involutions ................................ 33
   2.7 More on multiplication operators ............ 35
   2.8 Matrices ................................ 36
   2.9 Traces of matrices ........................... 37
   2.10 Vector spaces and linear mappings .......... 39
   2.11 Ideals and quotients ....................... 40
   2.12 Bilinear forms ............................ 42
   2.13 Dual spaces and mappings ................... 43
   2.14 Nondegenerate bilinear forms ............... 44
   2.15 Sesquilinear forms .......................... 46
3 Submultiplicativity and invertibility 49
  3.1 Invertibility ............................................. 49
  3.2 Submultiplicative seminorms .......................... 51
  3.3 Some matrix seminorms ................................. 52
  3.4 Continuity of inverses ................................. 54
  3.5 Banach algebras ....................................... 56
  3.6 Invertible linear mappings ........................... 58
  3.7 Isometric linear mappings ............................ 59
  3.8 Hilbert space isometries .............................. 61
  3.9 Preserving bilinear forms ............................. 62
  3.10 Preserving nondegenerate bilinear forms ............ 64
  3.11 Preserving sesquilinear forms ........................ 65
  3.12 Bilinear forms and matrices ........................ 67
  3.13 Sesquilinear forms and matrices .................... 70
  3.14 Invertibility and involutions ........................ 72
  3.15 Invertibility and seminorms .......................... 74

4 Formal power series 76
  4.1 Direct sums and products ................................ 76
  4.2 Bilinear mappings and Cauchy products ............... 78
  4.3 Formal power series and modules ...................... 80
  4.4 Sums and extensions .................................. 81
  4.5 Extending bilinear mappings ........................... 83
  4.6 Formal power series and algebras ..................... 85
  4.7 Invertibility in \(A[[T]]\) ................................. 86
  4.8 Homomorphisms over \(k[T]\) ............................. 87
  4.9 Formal power series and homomorphisms ............... 89
  4.10 Extensions and compositions .......................... 91
  4.11 Two-step extensions .................................. 93
  4.12 Extending algebra homomorphisms ..................... 94
  4.13 Some remarks about commutators ...................... 96
  4.14 Formal power series and involutions ................ 97

5 Some related notions 99
  5.1 Adjoining nilpotent elements .......................... 99
  5.2 Adjoining nilpotent elements to algebras ............. 100
  5.3 Invertibility in \(A[\epsilon_n]\) .......................... 102
  5.4 Adjoining two nilpotent elements ..................... 103
  5.5 Invertibility in \(A[\epsilon_1, \eta_1]\) ..................... 105
  5.6 Differentiation ....................................... 106
  5.7 Polynomial functions .................................. 108
  5.8 Several commuting indeterminates ..................... 109
  5.9 Polynomial functions in several variables ............ 112
  5.10 Partial derivatives ................................... 113
  5.11 Formal differential operators ......................... 116
  5.12 First-order differential operators ................... 117
5.13 Homogeneous formal polynomials ............................................ 119
5.14 Homogeneous differential operators ...................................... 120

6 Bilinear actions and representations ........................................ 122
6.1 Bilinear actions ..................................................................... 122
6.2 Subactions and homomorphisms ............................................ 123
6.3 Quotient actions .................................................................... 124
6.4 Representations of associative algebras ................................. 126
6.5 Representations of Lie algebras .............................................. 127
6.6 Subrepresentations ............................................................... 129
6.7 Homomorphisms between representations ............................. 130
6.8 Quotient representations ....................................................... 131
6.9 Sums of representations ....................................................... 132
6.10 Compatible bilinear mappings .............................................. 133
6.11 Representations and formal power series ............................. 135
6.12 Opposite algebras ............................................................... 136
6.13 Matrices and associative algebras ....................................... 138
6.14 Irreducibility ....................................................................... 140

7 Representations and multilinear mappings ................................. 141
7.1 Some remarks about subalgebras ........................................... 141
7.2 Representations on linear mappings ...................................... 142
7.3 Multilinear mappings ............................................................ 143
7.4 Boundedness and multilinearity ............................................ 145
7.5 Representations on multilinear mappings ............................. 147
7.6 Centralizers and invariant elements ....................................... 148
7.7 Invariant multilinear mappings ............................................. 149
7.8 Traces of linear mappings ..................................................... 151
7.9 The Killing form ................................................................... 153
7.10 Invariant subspaces and traces ............................................ 154
7.11 Radicals of bilinear mappings ............................................. 155
7.12 Tensor products ................................................................. 157
7.13 Functions on sets ............................................................... 158
7.14 Some natural isomorphisms ................................................ 159

8 Formal series and ordered rings .................................................. 161
8.1 Poles of finite order .............................................................. 161
8.2 Sequences and series ............................................................ 162
8.3 Formal series and module homomorphisms .......................... 164
8.4 Extending module homomorphisms ...................................... 165
8.5 Homomorphisms from \( A((T)) \) into \( B((T)) \) ......................... 166
8.6 Formal series and bilinear mappings ..................................... 168
8.7 Bilinear mappings on \( A((T)) \times B((T)) \) ............................. 170
8.8 Algebras and modules over \( k((T)) \) ....................................... 171
8.9 Absolute values on \( k((T)) \) .................................................. 173
8.10 Formal series and algebra homomorphisms .......................... 175
8.11 Involutions and formal series ........................................... 176
8.12 Ordered rings .............................................................. 178
8.13 Some additional features ................................................. 179
8.14 Ordered fields .............................................................. 180

9 Solvability and nilpotency ................................................. 182
  9.1 Some basic isomorphism theorems ...................................... 182
  9.2 Products of ideals ........................................................ 183
  9.3 Solvable Lie algebras ...................................................... 184
  9.4 The solvable radical ...................................................... 185
  9.5 Nilpotent Lie algebras .................................................... 187
  9.6 Two-dimensional Lie algebras ......................................... 188
  9.7 Nilpotency conditions .................................................... 189
  9.8 Maximal Lie subalgebras ............................................... 191
  9.9 Nilpotent linear mappings ............................................ 192
  9.10 Engel’s theorem .......................................................... 194
  9.11 Flags and matrices ..................................................... 195
  9.12 A useful lemma .......................................................... 197
  9.13 Lie’s theorem ............................................................ 199
  9.14 Structure constants ..................................................... 200
  9.15 Another corollary ....................................................... 201

10 Matrices and traces ....................................................... 203
  10.1 Some remarks about $gl_n(k)$ ......................................... 203
  10.2 Some basic properties of $sl_2(k)$ ................................... 205
  10.3 Scalar and diagonal matrices ........................................ 206
  10.4 Centrality in $gl_n(k), sl_n(k)$ ....................................... 207
  10.5 Solvability and traces ................................................ 208
  10.6 Diagonalizable linear mappings ..................................... 209
  10.7 Nilpotent vectors ....................................................... 211
  10.8 Jordan–Chevalley decompositions ................................... 212
  10.9 Some related situations ............................................... 214
  10.10 Replicas ................................................................. 215
  10.11 Two lemmas about traces ........................................... 217
  10.12 Cartan’s criterion ...................................................... 219
  10.13 Comparing radicals .................................................... 220
  10.14 Complementary ideals ................................................. 221
  10.15 Simple Lie algebras ................................................... 222

11 Some examples and related properties ................................ 224
  11.1 Simplicity and solvability ............................................ 224
  11.2 Traces on $M_n(k)$ ...................................................... 225
  11.3 A nice criterion ........................................................ 227
  11.4 Ideals and structure constants ....................................... 228
  11.5 Nondegeneracy and bilinear forms .................................. 229
  11.6 Bilinear forms and adjoints .......................................... 230
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.7</td>
<td>Bilinear forms and symmetry conditions</td>
<td>232</td>
</tr>
<tr>
<td>11.8</td>
<td>Traces and involutions</td>
<td>233</td>
</tr>
<tr>
<td>11.9</td>
<td>Inner products over ordered fields</td>
<td>234</td>
</tr>
<tr>
<td>11.10</td>
<td>Self-adjoint Lie subalgebras</td>
<td>237</td>
</tr>
<tr>
<td>11.11</td>
<td>Comparing involutions</td>
<td>238</td>
</tr>
<tr>
<td>11.12</td>
<td>Some projections</td>
<td>239</td>
</tr>
<tr>
<td>11.13</td>
<td>Antisymmetric bilinear forms</td>
<td>241</td>
</tr>
<tr>
<td>11.14</td>
<td>Symmetric bilinear forms</td>
<td>243</td>
</tr>
<tr>
<td>12.1</td>
<td>Complexifying ordered fields</td>
<td>245</td>
</tr>
<tr>
<td>12.2</td>
<td>Sesquilinearity over $k[\sqrt{-a}]$</td>
<td>246</td>
</tr>
<tr>
<td>12.3</td>
<td>Inner products over $k[\sqrt{-a}]$</td>
<td>248</td>
</tr>
<tr>
<td>12.4</td>
<td>Self-adjointness in $gl(V)$</td>
<td>250</td>
</tr>
<tr>
<td>12.5</td>
<td>Projections and sesquilinear forms</td>
<td>252</td>
</tr>
<tr>
<td>12.6</td>
<td>Nondegenerate Hermitian forms</td>
<td>253</td>
</tr>
<tr>
<td>12.7</td>
<td>Real parts</td>
<td>255</td>
</tr>
<tr>
<td>12.8</td>
<td>Bilinear forms over $k[\sqrt{-a}]$</td>
<td>257</td>
</tr>
<tr>
<td>12.9</td>
<td>Conjugate-linearity and adjoints</td>
<td>258</td>
</tr>
<tr>
<td>12.10</td>
<td>Antisymmetric forms over $k[\sqrt{-a}]$</td>
<td>259</td>
</tr>
<tr>
<td>12.11</td>
<td>Symmetric forms over $k[\sqrt{-a}]$</td>
<td>261</td>
</tr>
<tr>
<td>12.12</td>
<td>Nonnegative self-adjoint operators</td>
<td>262</td>
</tr>
<tr>
<td>12.13</td>
<td>Polar decompositions</td>
<td>263</td>
</tr>
<tr>
<td>12.14</td>
<td>Bilinear forms and inner products</td>
<td>265</td>
</tr>
<tr>
<td>12.15</td>
<td>Complex bilinear forms</td>
<td>266</td>
</tr>
<tr>
<td>13.1</td>
<td>Complete reducibility</td>
<td>269</td>
</tr>
<tr>
<td>13.2</td>
<td>Casimir elements</td>
<td>270</td>
</tr>
<tr>
<td>13.3</td>
<td>Another perspective</td>
<td>272</td>
</tr>
<tr>
<td>13.4</td>
<td>A more particular situation</td>
<td>274</td>
</tr>
<tr>
<td>13.5</td>
<td>A splitting theorem</td>
<td>274</td>
</tr>
<tr>
<td>13.6</td>
<td>Weyl’s theorem</td>
<td>276</td>
</tr>
<tr>
<td>13.7</td>
<td>Symmetric forms and tensors</td>
<td>277</td>
</tr>
<tr>
<td>13.8</td>
<td>Reductive Lie algebras</td>
<td>279</td>
</tr>
<tr>
<td>13.9</td>
<td>Semisimple ideals</td>
<td>279</td>
</tr>
<tr>
<td>13.10</td>
<td>Another approach</td>
<td>280</td>
</tr>
<tr>
<td>13.11</td>
<td>Inner derivations</td>
<td>282</td>
</tr>
<tr>
<td>14.1</td>
<td>A duality argument</td>
<td>283</td>
</tr>
<tr>
<td>14.2</td>
<td>A criterion for nilpotency</td>
<td>284</td>
</tr>
<tr>
<td>14.3</td>
<td>ad-Diagonalizability</td>
<td>285</td>
</tr>
<tr>
<td>14.4</td>
<td>Semisimple subalgebras of $gl(V)$</td>
<td>287</td>
</tr>
<tr>
<td>14.5</td>
<td>Tensors</td>
<td>288</td>
</tr>
<tr>
<td>14.6</td>
<td>Induced linear mappings</td>
<td>289</td>
</tr>
</tbody>
</table>
## CONTENTS

14.7 Invariant tensors ........................................... 291
14.8 Diagonalizability and quotients ............................. 291
14.9 Diagonalizability and representations ...................... 293
14.10 Exponentiating nilpotents .................................. 294
14.11 Nilpotent derivations ...................................... 296
14.12 Exponentials and homomorphisms .......................... 298
14.13 Representations and structure constants ................... 299
14.14 Lie's theorem and nilpotency ............................... 300

15 Complexifications and $sl_2(k)$ modules ....................... 302
  15.1 Basic properties of $sl_2(k)$ modules ...................... 302
  15.2 Maximal or primitive vectors ............................... 303
  15.3 Related submodules ....................................... 304
  15.4 Constructing modules $W(m)$ ............................... 306
  15.5 More on $W(m)$ ........................................... 307
  15.6 Another construction ..................................... 309
  15.7 Some exponentials ....................................... 310
  15.8 Exponentials and $W(m)$ .................................. 312
  15.9 Additional properties of $sl_2(k)$ modules ................ 313
  15.10 Complexifying real vector spaces .......................... 314
  15.11 Spaces of linear mappings ................................ 315
  15.12 Complexifying algebras over $\mathbb{R}$ .................... 317
  15.13 Conjugate-linear involutions ............................. 318

16 Module homomorphisms ........................................ 321
  16.1 Bilinear actions and homomorphisms ....................... 321
  16.2 Homomorphisms and semisimplicity ........................ 322
  16.3 Linear mappings and dimensions ........................... 324
  16.4 Homomorphisms and dimensions ........................... 324
  16.5 Semisimplicity and dimensions ............................ 325
  16.6 Absolutely simple Lie algebras ............................ 326
  16.7 Algebras over subrings .................................... 328
  16.8 Lie algebras over subfields ................................ 329

17 Subalgebras and diagonalizability .............................. 331
  17.1 Toral subalgebras ....................................... 331
  17.2 Simultaneous diagonalizability ............................ 332
  17.3 Related bilinear forms ................................... 334
  17.4 Maximal toral subalgebras ................................ 335
  17.5 Self-centralizability and diagonalizability ............... 337
  17.6 Characteristic 0, $Z(A) = \{0\}$ .......................... 339
  17.7 The dimension of $A_\alpha$ ................................. 340
  17.8 Other elements of $\Phi_B$ ................................ 343
  17.9 Adding elements of $\Phi_B$ ............................... 345

Bibliography .................................................... 348
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,

$$d(x, x) = 0$$

(1.1.1) for every $x \in X$. Second,

$$d(x, y) = d(y, x)$$

(1.1.2) for every $x, y \in X$. Third,

$$d(x, z) \leq d(x, y) + d(y, z)$$

(1.1.3) for every $x, y, z \in X$. If we also have that

$$d(x, y) > 0$$

(1.1.4) 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 it satisfies (1.1.1), (1.1.2), and

$$d(x, z) \leq \max(d(x, y), d(y, z))$$

(1.1.5) 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

$$B(x, r) = B_d(x, r) = \{y \in X : d(x, y) < r\}.$$
Similarly, the closed ball in $X$ centered at $x \in X$ with radius $r \geq 0$ is defined by
\[(1.1.7) \quad \overline{B}(x, r) = \overline{B}_d(x, r) = \{y \in X : d(x, y) \leq r\}.
\]
A subset $U$ 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
\[(1.1.8) \quad B(x, r) \subseteq U.
\]
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
\[(1.1.9) \quad B(x, r) \subseteq B(y, r).
\]
More precisely,
\[(1.1.10) \quad B(x, r) = B(y, r),
\]
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
\[(1.1.11) \quad \overline{B}(x, r) \subseteq \overline{B}(y, r),
\]
and hence
\[(1.1.12) \quad \overline{B}(x, r) = \overline{B}(y, r).
\]
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,
\[(1.2.1) \quad |xy| = |x||y|
\]
for every $x, y \in k$. Third,
\[(1.2.2) \quad |x + y| \leq |x| + |y|
\]
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
1.2. ABSOLUTE VALUE FUNCTIONS

(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

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

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

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

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

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

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

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

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

\[
|n \cdot 1| \leq 1
\]

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

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

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

\[
r + t = r^{1-a} r^a + t^{1-a} t^a \leq \max(r^{1-a}, t^{1-a}) (r^a + t^a).
\]

If \( a \leq 1 \), then it follows that

\[
r + t \leq \max(r, t)^{1-a} (r^a + t^a) \leq (r^a + t^a)^{(1-a)/a+1} = (r^a + t^a)^{1/a},
\]

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.
Let $|\cdot|_1$ and $|\cdot|_2$ be absolute value functions on $k$. If there is a positive real number $a$ such that
\begin{equation}
|x|_2 = |x|_1^a
\end{equation}
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 $\mathbb{Q}_p$. If $x \in \mathbb{Q}_p$ 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 $\mathbb{R}$ or $\mathbb{C}$, in such a way that $| \cdot |$ corresponds to an absolute value function on $\mathbb{R}$ or $\mathbb{C}$ 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

$$
\{|x| : x \in k, x \neq 0\}
$$

is a subgroup of the group $\mathbb{R}_+$ of positive real numbers with respect to multiplication. Of course, (1.5.1) is the trivial subgroup $\{1\}$ of $\mathbb{R}_+$ 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 $\mathbb{R}$, then $| \cdot |$ is said to be discrete on $k$.

Put

$$
\rho_1 = \sup\{|x| : x \in k, |x| < 1\},
$$

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 $\mathbb{R}$ 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 $\mathbb{Q}$ into $k$. This leads to an absolute value function on $\mathbb{Q}$, using $| \cdot |$ on $k$. It is easy to see that $\mathbb{Q}$ 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 $\mathbb{Q}$ mentioned in Section 1.3, we get that this absolute value function on $\mathbb{Q}$ is equivalent to the standard absolute value function. In particular, it follows that this absolute value function on $\mathbb{Q}$ 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 $k$. 

1.6. *P*-adic integers

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

\[(1.5.3) \quad |y| \leq \rho_1 |z|,\]

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 $\rho_1$, 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

\[(1.5.4) \quad |x - y| < |y|,\]

then

\[(1.5.5) \quad |x| = |y|.\]

More precisely,

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

by the ultrametric version of the triangle inequality. Similarly,

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

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

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

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

\[(1.6.1) \quad (1 - x) \sum_{j=0}^{n} x^j = \sum_{j=0}^{n} x^j - \sum_{j=0}^{n} x^{j+1} = \sum_{j=0}^{n} x^j - \sum_{j=1}^{n+1} x^j = 1 - x^{n+1},\]

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

\[(1.6.2) \quad \sum_{j=0}^{n} x^j = (1 - x^{n+1}) (1 - x)^{-1}.\]
Let \(| \cdot |\) be an absolute value function on \(k\), so that

\[
| \sum_{j=0}^{n} x^j - (1 - x)^{-1} | = |x|^{n+1} (1 - x)^{-1} = |x|x^{n+1} |1 - x|^{-1}.
\]  

If \(|x| < 1\), then we get that

\[
\lim_{n \to \infty} \left| \sum_{j=0}^{n} x^j - (1 - x)^{-1} \right| = 0.
\]  

Let \(p\) be a prime number, and let \(y\) be an integer. Thus \(x = p\ y\) satisfies

\[
|x|_p = p^{-1} |y|_p \leq p^{-1} < 1,
\]  

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

\[
\lim_{n \to \infty} \left| \sum_{j=0}^{n} p^j y^j - (1 - p y)^{-1} \right|_p = 0,
\]  

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 \(b c \equiv 1 \mod p\), because the integers modulo \(p\) form a field. Thus \(z\) can be expressed as

\[
z = (a c)/(b c) = a c (1 - p y)^{-1},
\]  

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

\[
\mathbb{Z}_p = \{ x \in \mathbb{Q}_p : |x|_p \leq 1 \},
\]

which is the set of \(p\)-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 \(\mathbb{Q}_p\) 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 \(\mathbb{Q}_p\) 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 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 $\overline{B}(0, 1)/\overline{B}(0, 1)$ (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).
is defined as a commutative ring when $j \geq 0$. There is a natural ring homomorphism from $\mathbb{Z}$ into (1.7.5), which is the composition of the inclusion of $\mathbb{Z}$ in $\mathbb{Z}_p$ with the quotient mapping from $\mathbb{Z}_p$ onto (1.7.5). Observe that

\[(1.7.6) \quad \mathbb{Z} \cap (p^j \mathbb{Z}_p) = p^j \mathbb{Z}\]

for every nonnegative integer $j$, which is the kernel of the homomorphism from $\mathbb{Z}$ into (1.7.5) just mentioned. Thus we get an injective ring homomorphism from

\[(1.7.7) \quad \mathbb{Z}/(p^j \mathbb{Z})\]

into (1.7.6) for each nonnegative integer $j$. One can check that this homomorphism is surjective, because $\mathbb{Z}$ is dense in $\mathbb{Z}_p$, as in the previous section. This shows that (1.7.5) is isomorphic to (1.7.7) as a ring for every nonnegative integer $j$.

### 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,

\[(1.8.1) \quad N(tv) = |t|N(v)\]

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

\[(1.8.2) \quad N(v + w) \leq N(v) + N(w)\]

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

\[(1.8.3) \quad N(v) > 0\]

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

\[(1.8.4) \quad N(v + w) \leq \max(N(v), N(w))\]

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

$$d(v, w) = d_N(v, w) = N(v - w)$$  \hspace{1cm} (1.8.5)$$
defines a semimetric on $V$. If $N$ is a norm on $V$, then (1.8.5) is a metric on $V$. Thus (1.8.5) is an ultrametric on $V$ when $N$ is an ultranorm on $V$.

Suppose for the moment that $| \leq |$ 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 $| \leq |$ 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 $N$. If $N$ is a norm on $V$ with respect to $| \cdot |$ on $k$, then $N(v)^a$ is a norm on $V$ with respect to $| \cdot |^a$ on $k$.

If $| \leq |$ 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

$$N(v - w) < N(w),$$  \hspace{1cm} (1.8.6)$$
then

$$N(v) = N(w).$$  \hspace{1cm} (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

$$N(v) \leq \max(N(v - w), N(w)) = N(w),$$  \hspace{1cm} (1.8.8)$$
by the semi-ultranorm version of the triangle inequality. We also have that

$$N(w) \leq \max(N(v - w), N(v)),$$  \hspace{1cm} (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 $N$
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 semimetrics 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.

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 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

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

for every $v, v' \in V$, and in particular that $T$ is continuous with respect to the semimetrics 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 semimetrics, 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 $\mathcal{BL}(V, W)$ be the space of bounded linear mappings from $V$ into $W$, with respect to $N_V$ and $N_W$. If $T \in \mathcal{BL}(V, W)$, then put

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

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 $C = \|T\|_{\text{op}}$. One can verify that $\mathcal{BL}(V, W)$ 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 $N_W$ is a norm on $W$, then (1.9.3) is a norm on $\mathcal{BL}(V, W)$. If $N_W$ is a semi-ultranorm on $W$, then (1.9.3) is a semi-ultranorm on $\mathcal{BL}(V, W)$. In particular, if $N_W$ is an ultranorm on $W$, then (1.9.3) is an ultranorm on $\mathcal{BL}(V, W)$.

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

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

\[(1.9.5) \quad \|T_2 \circ T_1\|_{\text{op}, VZ} \leq \|T_1\|_{\text{op}, VW} \|T_2\|_{\text{op}, WZ}.\]

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 $\mathcal{B}\mathcal{L}(V, W)$ 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 $\mathcal{B}\mathcal{L}(V, W)$ 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 $k^n$

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

\[(1.10.1) \quad \|v\|_1 = \sum_{j=1}^{n} |v_j|\]

and

\[(1.10.2) \quad \|v\|_\infty = \max_{1 \leq j \leq n} |v_j|\]
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
\begin{equation}
\|v\|_\infty \leq \|v\|_1 \leq n \|v\|_\infty
\end{equation}
for every $v \in k^n$. Let
\begin{equation}
 d_1(v,w) = \|v - w\|_1
\end{equation}
and
\begin{equation}
 d_\infty(v,w) = \|v - w\|_\infty
\end{equation}
be the metrics on $k^n$ associated to (1.10.1) and (1.10.2), respectively. Thus
\begin{equation}
 d_\infty(v,w) \leq d_1(v,w) \leq n d_\infty(v,w)
\end{equation}
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
\begin{equation}
 v = \sum_{l=1}^n v_l e_l
\end{equation}
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
\begin{equation}
 T(v) = T\left(\sum_{l=1}^n v_l e_l\right) = \sum_{l=1}^n v_l T(e_l)
\end{equation}
for every $v \in k^n$, and hence
\begin{equation}
 N_W(T(v)) \leq \sum_{l=1}^n |v_l| N_W(T(e_l)).
\end{equation}
In particular,
\begin{equation}
 N_W(T(v)) \leq \left(\max_{1 \leq l \leq n} N_W(T(e_l))\right) \|v\|_1
\end{equation}
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
\begin{equation}
 \max_{1 \leq l \leq n} N_W(T(e_l)).
\end{equation}
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,
1.11. INNER PRODUCTS

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

$$\sum_{l=1}^{n} N_{W}(T(e_l)).$$

(1.10.13)

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

$$(1.10.14) N_{W}(T(v)) \leq \max_{1 \leq l \leq n} (|v_l| N_{W}(T(e_l))) \leq \left( \max_{1 \leq l \leq n} N_{W}(T(e_l)) \right) \|v\|_\infty$$

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 $C$ of complex numbers, so that $V$ and $W$ may be considered as vector spaces over $R$ 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 $R$, and that $T$ is complex-linear if $T$ is linear as a mapping from $V$ into $W$ as vector spaces over $C$. Thus a complex-linear mapping $T$ from $V$ into $W$ is the same as a real-linear mapping that also satisfies

$$(1.11.1) T(iv) = iT(v)$$

for every $v \in V$. A real-linear mapping $T$ from $V$ into $W$ is said to be conjugate-linear if

$$(1.11.2) T(iv) = -iT(v)$$

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

$$(1.11.3) T(\overline{a}v) = \overline{a}T(v)$$

for every $a \in C$ and $v \in V$, where $\overline{a}$ is the usual complex-conjugate of $a$.

Suppose from now on in this section that $k = R$ or $C$, 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

$$(1.11.4) \langle w, v \rangle = \langle v, w \rangle$$
for every $v, w \in V$ in the real case, and that

\[
\langle w, v \rangle = \langle v, w \rangle
\]

(1.11.5)

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

\[
\langle v, v \rangle > 0
\]

(1.11.6)

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

\[
\|v\| = \langle v, v \rangle^{1/2}
\]

(1.11.7)

for every $v \in V$, using the nonnegative square root on the right side. It is well known that

\[
\langle v, w \rangle \leq \|v\| \|w\|
\]

(1.11.8)

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

\[
\langle v, w \rangle = \langle v, w \rangle_{\mathbb{R}^n} = \sum_{j=1}^{n} v_j w_j.
\]

(1.11.9)

Similarly, the standard inner product on $\mathbb{C}^n$ is given by

\[
\langle v, w \rangle = \langle v, w \rangle_{\mathbb{C}^n} = \sum_{j=1}^{n} v_j \overline{w_j}.
\]

(1.11.10)

In both cases, the corresponding norm is given by

\[
\|v\| = \|v\|_2 = \left( \sum_{j=1}^{n} |v_j|^2 \right)^{1/2}.
\]

(1.11.11)

It is easy to see that

\[
\|v\|_\infty \leq \|v\|_2 \leq n^{1/2} \|v\|_\infty
\]

(1.11.12)

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

\[
\|v\|_2 \leq \|v\|_1 \leq n^{1/2} \|v\|_2
\]

(1.11.13)

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
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

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

(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 \(\mathcal{BL}(V,W)\) into \(\mathcal{BL}(W,V)\) 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).
1.13. BOUNDED BILINEAR MAPPINGS

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
\begin{equation}
\langle v_j, v_l \rangle = 0
\end{equation}
when $j \neq l$. This implies that
\begin{equation}
\left\| \sum_{j=1}^{n} v_j \right\|^2 = \sum_{j=l}^{n} \| v_j \|^2
\end{equation}
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
\begin{equation}
\left\| \sum_{j=1}^{\infty} v_j \right\|^2 = \sum_{j=1}^{\infty} \| v_j \|^2.
\end{equation}
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
\begin{equation}
N_Z(b(v, w)) \leq C N_V(v) N_W(w)
\end{equation}
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
\begin{equation}
b(v, w) - b(v', w') = b(v - v', w) + b(v', w - w'),
\end{equation}
so that
\begin{equation}
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.
\end{equation}
One can use this to check that $b$ is continuous with respect to the seminorms associated to $N_V$, $N_W$, $N_Z$ and the corresponding product topology on $V \times W$. 
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 $V_0 \times W_0$ 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 $e_1^V, \ldots, e_{n_V}^V$ and $e_1^W, \ldots, e_{n_W}^W$ 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^V, e_l^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^V, e_l^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^V, e_l^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}
as in (1.10.2), then (1.13.5) implies that (1.13.1) holds, with $C$ equal to

\[(1.13.11) \quad \sum_{j=1}^{n_V} \sum_{l=1}^{n_W} N_Z(b(e_j^{n_V}, e_l^{n_W})).\]

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

\[(1.13.12) \quad N_Z(b(v, w)) \leq \max\{|v_j| \cdot |w_l| \cdot N(b(e_j^{n_V}, e_l^{n_W})): 1 \leq j \leq n_V, 1 \leq l \leq n_W\}\]

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

\[(1.14.1) \quad t E = \{tv: v \in E\}.\]

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

\[(1.14.2) \quad t E \subseteq E\]

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

\[(1.14.3) \quad t E = E,\]

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

\[(1.14.4) \quad v \in t_1 A.\]

This implies that

\[(1.14.5) \quad v \in t A\]

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

\[(1.14.6) \quad t^{-1} v \in A\]

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
\[ N(tv) = |t|N(v) \tag{1.14.7} \]
for every $t \in k$ and $v \in V$. Put
\[ B_N(0, r) = \{ v \in V : N(v) < r \} \tag{1.14.8} \]
for every positive real number $r$, and
\[ B(0, r) = \{ v \in V : N(v) \leq r \} \tag{1.14.9} \]
for every nonnegative real number $r$. If $t \in k$ and $t \neq 0$, then
\[ tB_N(0, r) = B_N(0, |t|r) \tag{1.14.10} \]
for every $r > 0$, and
\[ tB_N(0, r) = B_N(0, |t| r) \tag{1.14.11} \]
for every $r \geq 0$. In particular, $B_N(0, r)$ is balanced in $V$ for every $r > 0$, and $B_N(0, r)$ 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 $B_N(0, r)$ 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
\[ N_A(v) = \inf \{ |t| : t \in k, t \neq 0, v \in tA \} \tag{1.14.12} \]
\[ = \inf \{ |t| : t \in k, t \neq 0, t^{-1}v \in A \} \]
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
\[ N_A(t'v) = |t'|N_A(v) \tag{1.14.13} \]
for every $t' \in k$ and $v \in V$. If $v \in A$, then $N_A(v) \leq 1$, so that
\[ A \subseteq \overline{B}_N(0, 1). \tag{1.14.14} \]
If $v \in V$ satisfies $N_A(v) < 1$, then there is a $t \in k$ such that $|t| < 1$ and $v \in tA$. This implies that $v \in A$, because $A$ is balanced in $V$, so that
\[ A = \overline{B}_N(0, 1) \tag{1.14.15} \]
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 $v \in tA$, 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
\[ A = \overline{B}_N(0, 1) \tag{1.14.16} \]
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 $B_N(0, r)$ 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 $B_N(0, r)$ is a balanced subgroup in $V$ for every $r > 0$, and that $\overline{B}_N(0, r)$ 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 $tv \in E$ 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

$$N_A(v), N_A(w) < r.$$ (1.15.2)

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

$$|t_1(v)|, |t_1(w)| < r$$ (1.15.3)

and

$$v \in t_1(v) A, w \in t_1(w) A,$$ (1.15.4)
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}

Thus $t_1(v)A, t_1(w)A \subseteq t_1A$, because $A$ is balanced. This implies that $v$ and $w$ are both elements of $t_1A$, by (1.15.4). It follows that

\begin{equation}
v + w \in t_1A,
\end{equation}

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}

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}

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
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

\begin{equation}
\beta(b, a) = \beta(a, b) \tag{2.1.1}
\end{equation}

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

\begin{equation}
\beta(b, a) = -\beta(a, b) \tag{2.1.2}
\end{equation}

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

\begin{equation}
\beta(a, a) = 0 \tag{2.1.3}
\end{equation}

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

\begin{equation}
\beta(a + b, a + b) = \beta(a, a) + \beta(a, b) + \beta(b, a) + \beta(b, b) \tag{2.1.4}
\end{equation}

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

\begin{equation}
(a, b) \mapsto ab, \tag{2.2.1}
\end{equation}

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

\begin{equation}
\beta(b, a) = \beta(a, b) \tag{2.1.1}
\end{equation}

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

\begin{equation}
\beta(b, a) = -\beta(a, b) \tag{2.1.2}
\end{equation}

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

\begin{equation}
\beta(a, a) = 0 \tag{2.1.3}
\end{equation}

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

\begin{equation}
\beta(a + b, a + b) = \beta(a, a) + \beta(a, b) + \beta(b, a) + \beta(b, b) \tag{2.1.4}
\end{equation}

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

\begin{equation}
(a, b) \mapsto ab, \tag{2.2.1}
\end{equation}
and we may refer to $a\ b$ as the product of $a, b \in A$. If this bilinear mapping is symmetric, so that

\[(2.2.2)\quad a\ b = b\ a\]

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

\[(2.2.3)\quad (a\ b)\ c = a\ (b\ c)\]

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

\[(2.2.4)\quad e\ a = a\ e = a\]

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 $a\ b \in A_0$, 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

\[(2.2.5)\quad \phi(a_1\ a_2) = \phi(a_1)\ \phi(a_2)\]

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

\[(2.2.7)\quad M_a(x) = a\ x\]
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

\[
a \mapsto M_a
\]

(2.2.8)

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

\[
M_a(e) = a e = a
\]

(2.2.9)

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

If \( a, b, x \in A \), then

\[
M_a(M_b(x)) = M_a(b x) = a (b x)
\]

(2.2.10)

and

\[
M_{a b}(x) = (a b) x.
\]

(2.2.11)

If \( A \) is an associative algebra, then it follows that

\[
M_a \circ M_b = M_{a b},
\]

(2.2.12)

as mappings from \( A \) into itself. This means that (2.2.8) is an algebra homomorphism from \( A \) into \( \text{Hom}_k(A, A) \), using composition as multiplication on the space \( \text{Hom}_k(A, A) \) of module homomorphisms from \( A \) into itself, as before.

### 2.3 Lie algebras

Let \( k \) be a commutative ring with a multiplicative identity element again, and let \( A \) be a module over \( k \). Also let \([x, y]\) be a mapping from \( A \times A \) into \( A \) that is bilinear over \( k \). Suppose that

\[
[x, x] = 0
\]

(2.3.1)

for every \( x \in A \). This implies that

\[
[y, x] = -[x, y]
\]

(2.3.2)

for every \( x, y \in A \), as in Section 2.1. If \( 1 + 1 \) has a multiplicative inverse in \( k \), then (2.3.2) implies (2.3.1), as before.

The Jacobi identity may be formulated as saying that

\[
[[x, y], z] + [[y, z], x] + [[z, x], y] = 0
\]

(2.3.3)
for every \( x, y, z \in A \). Alternatively, the Jacobi identity may be expressed as saying that
\[(2.3.4) \quad [x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0\]
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 \textit{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
\[(2.3.5) \quad [x, y] = 0\]
for every \( x, y \in A \), then \( A \) is said to be \textit{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
\[(2.3.6) \quad [x, y] = xy - yx\]
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 \((A, [\cdot, \cdot])\) be a Lie algebra over \( k \), and let \( A_0 \) be a submodule of \( A \), as a module over \( k \). As in the previous section, \( A_0 \) is said to be a \textit{subalgebra of} \( A \) if \([x, y] \in A_0 \) for every \( x, y \in A_0 \). In this situation, \( A_0 \) is a Lie algebra over \( k \) with respect to the restriction of \([x, y]\) to \( x, y \in A_0 \). One may also refer to \( A_0 \) as a \textit{Lie subalgebra} of \( A \). If \( A \) is an associative algebra over \( k \), and \( A_0 \) is a subalgebra of \( A \), then \( A_0 \) is also a Lie subalgebra of \( A \) as 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
\[(2.3.7) \quad \phi([x, y]_A) = [\phi(x), \phi(y)]_B\]
for every \( x, y \in A \), then \( \phi \) is said to be a \textit{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 $[\cdot, \cdot]_A$ and $[\cdot, \cdot]_B$ be the corresponding commutators on $A$ and $B$, as in (2.3.6). If $x, y \in A$, then

\begin{equation}
(2.3.8) \quad \phi([x, y]_A) = \phi(x y - y x) = \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 $[\cdot, \cdot]_A$ and $[\cdot, \cdot]_B$ 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 $[\cdot, \cdot]_A$ and $[\cdot, \cdot]_B$, 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 ad $x$ be the mapping from $A$ into itself defined by

\begin{equation}
(2.4.1) \quad (\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, ad $x$ is a module homomorphism from $A$ into itself, as a module over $k$, because $[x, y]_A$ is linear over $k$ in $y$. Thus

\begin{equation}
(2.4.2) \quad 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}
(2.4.3) \quad [\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}
(2.4.4) \quad ([\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
\[
(2.4.5) \quad (\text{ad}[x, y]_A)(z) = [[x, y]_A, z]_A = -[z, [x, y]_A],
\]
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
\[
(2.4.6) \quad (\text{ad}[x, y]_A)(z) = (\text{ad} x, \text{ad} y)(z)
\]
for every \(z \in A\), so that
\[
(2.4.7) \quad \text{ad}[x, y]_A = [\text{ad} x, \text{ad} y]
\]
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 \(a \cdot b\). Also let \(\delta\) be a module homomorphism from \(A\) into itself, as a module over \(k\). If \(\delta\) satisfies the product rule
\[
(2.5.1) \quad \delta(a \cdot b) = \delta(a) \cdot b + a \cdot \delta(b)
\]
for every \(a, b \in A\), then \(\delta\) is said to be a derivation on \(A\). Let \(\text{Der}(A)\) be the collection of derivations on \(A\). Remember that the space \(\text{Hom}_k(A, A)\) 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 \(\text{Der}(A)\) is a submodule of \(\text{Hom}_k(A, A)\), as a module over \(k\). We also have that \(\text{Hom}_k(A, A)\) is an associative algebra over \(k\), with respect to composition of mappings. This implies that \(\text{Hom}_k(A, A)\) is a Lie algebra over \(k\) with respect to the corresponding commutator bracket. It is well known that \(\text{Der}(A)\) is a Lie subalgebra of \(\text{Hom}_k(A, A)\) with respect to the commutator bracket.

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

Let \( a \in A \) be given, and put
\[
\delta_a(x) = ax - xa
\]
(2.5.5)
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,
\[
a \mapsto \delta_a
\]
(2.5.6)
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
\[
\delta_a(x y) = ax y - (xy)a = \delta_a(x) y + x \delta_a(y),
\]
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
\[
\delta([a, b]_A) = \delta(a) b + [a, \delta(b)]_A
\]
(2.5.8)
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 \( \text{ad} \ x = [x, \cdot, \cdot]_A \in \text{Der}(A) \). If \( y, z \in A \), then
\[
(\text{ad} \ x)([y, z]_A) = [x, [y, z]_A]_A = -[y, [z, x]_A]_A - [z, [x, y]_A],
\]
(2.5.9)
using the Jacobi identity in the second step. It follows that
\[
(\text{ad} \ x)([y, z]_A) = [[x, y]_A, z]_A + [y, [x, z]_A],
\]
(2.5.10)
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 \( [x, y] = xy - yx \) be the usual commutator bracket corresponding to multiplication in \( A \). Let \( \delta \in \text{Der}(A) \) be given, and observe that
\[
\delta([x, y]) = \delta(xy - yx) = \delta(x)y + x\delta(y) - \delta(y)x - y\delta(x) = [\delta(x), y] + [x, \delta(y)]
\]
(2.5.11)
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 $x y$. 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

$$\phi((x^*)^*) = x$$

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
\begin{equation}
(2.6.7) \quad a + a^*
\end{equation}
is self-adjoint with respect to $x \mapsto x^*$, and
\begin{equation}
(2.6.8) \quad a - a^*
\end{equation}
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
\begin{equation}
(2.6.9) \quad ([a, b])^* = -[a^*, b^*] = -[-a, -b] = -[a, b],
\end{equation}
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 $ab$. If $a \in A$, then

\begin{equation}
M_a(x) = ax
\end{equation}

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,

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

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,

\begin{equation}
a \mapsto \tilde{M}_a
\end{equation}

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 $\tilde{M}_e$ is the identity mapping on $A$. We also have that

\begin{equation}
\tilde{M}_a(e) = ea = a
\end{equation}

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

Observe that

\begin{equation}
\tilde{M}_a(\tilde{M}_b(x)) = \tilde{M}_a(xb) = (xb)a
\end{equation}

and

\begin{equation}
\tilde{M}_{ba}(x) = x(ba)
\end{equation}

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

\begin{equation}
\tilde{M}_a \circ \tilde{M}_b = \tilde{M}_{ab}
\end{equation}

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

\begin{equation}
\tilde{M}_{[a, b]} = \tilde{M}_{ab - ba} = \tilde{M}_{a b} - \tilde{M}_{b a} = \tilde{M}_b \circ \tilde{M}_a - \tilde{M}_a \circ \tilde{M}_b
\end{equation}

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

\begin{equation}
M_{[a, b]} = M_{ab - ba} = M_{a b} - M_{b a} = M_a \circ M_b - M_b \circ M_a
\end{equation}
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(	ilde{M}_b(x)) = M_a(x b) = a(x b) \tag{2.7.10}
\end{equation}
and
\begin{equation}
\tilde{M}_b(M_a(x)) = \tilde{M}_b(a x) = (a x) b. \tag{2.7.11}
\end{equation}
These are the same when $A$ is an associative algebra, in which case
\begin{equation}
M_a \circ \tilde{M}_b = \tilde{M}_b \circ M_a \tag{2.7.12}
\end{equation}
for every $a, b \in A$.

If $a \in A$, then let $\text{ad} a$ be the mapping from $A$ into itself defined by
\begin{equation}
(\text{ad} a)(x) = [a, x] = a x - x a = M_a(x) - \tilde{M}_a(x) \tag{2.7.13}
\end{equation}
for every $x \in A$. Equivalently,
\begin{equation}
\text{ad} a = M_a - \tilde{M}_a, \tag{2.7.14}
\end{equation}
which is a module homomorphism from $A$ into itself, as a module over $k$. We also have that
\begin{equation}
a \mapsto \text{ad} a \tag{2.7.15}
\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}
\text{ad}[a, b] = [\text{ad} a, \text{ad} b] = (\text{ad} a) \circ (\text{ad} b) - (\text{ad} b) \circ (\text{ad} a) \tag{2.7.16}
\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 $\tilde{M}_a$.

### 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 $x y$. If $a, b \in M_n(A)$, then their product $c = a b$ is defined as usual by
\begin{equation}
c_{j,r} = \sum_{l=1}^{n} a_{j,l} b_{l,r} \tag{2.8.1}
\end{equation}
2.9. TRACES OF MATRICES

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 $a^t = (a^t_{l,j}) \in M_n(A)$ is defined as usual by

$$a^t_{j,l} = a_{l,j} \quad (2.8.2)$$

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

$$ (a^*)_{j,l} = (a_{l,j})^* \quad (2.8.3) $$

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

$$ (a^*)_{j,l} = \overline{a_{l,j}} \quad (2.8.4) $$

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.

2.9 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
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 \( x y \). If \( a, b \in M_n(A) \), then the products \( a b \) and \( b a \) are defined as elements of \( M_n(A) \) as in the previous section. Observe that

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

(2.9.3)

and

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

(2.9.4)

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

\[
\text{tr}(a b) = \text{tr}(b a)
\]

(2.9.5)

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

\[
\text{tr}(a b - b a) = 0
\]

(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 \([a, b] = a b - b a\) as the bilinear operation on \( gl_n(A) \), where the products \( a b \) and \( b a \) 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

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

(2.9.7)

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

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

(2.9.8)

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 \([a, b]\).
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 to composition of mappings.

The general linear algebra \( \mathfrak{gl}(V) \) 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 \( k \), \( \mathfrak{gl}(V) \) is the same as \( \mathcal{L}(V) \). If \( T_1 \) and \( T_2 \) are linear mappings from \( V \) into itself, then their commutator

\[
[T_1,T_2] = T_1 \circ T_2 - T_2 \circ T_1
\]

defines a linear mapping from \( V \) into itself as well. This defines a bilinear operation on \( \mathfrak{gl}(V) \), which we use to define the Lie bracket on \( \mathfrak{gl}(V) \). 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

\[
v = \sum_{l=1}^{n} t_l v_l,
\]

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

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

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

\[
a \mapsto T_a
\]

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

\[
T_a \circ T_b = T_{ab}
\]

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.
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

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

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

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

by (2.10.9). In particular, \( \mathfrak{sl}(V) \) 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 \( \mathfrak{sl}_n(k) \) onto \( \mathfrak{sl}(V) \).

### 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 \( A \), 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 \( a \cdot b \). Also let \( A_0 \) be a submodule of \( A \), as a module over \( k \). If

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

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 \textit{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

$$(a, b) \mapsto q_0(ab)$$

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

$q_0(ab) = 0$

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 $q_0(ab)$ 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 $q_0(ab)$ 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

$$(a, b)_A \in A_0$$

for every $a \in A$ and $b \in A_0$, then $A_0$ is said to be an \textit{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 $ab$. 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 $[a, b]_A = ab - ba$. Under these conditions, $A_0$ may be considered as 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

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

(2.12.1)

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

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

(2.12.2)

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

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

(2.12.3)

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

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

(2.12.4)

is symmetric on \( A \times A \), and

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

(2.12.5)

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

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

(2.12.6)

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

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

(2.12.7)

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 \( \text{Hom}_k(A, A) \). 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

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

(2.12.8)

for every $a, b \in A$. This also works when $\phi$ and $\psi$ are both antisymmetric with respect to $\beta$ on $A$, 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 $\theta$ is a linear functional on $W$, then $T' = \theta \circ T$ 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

$$T \mapsto T'$$

(2.13.2)

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

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

(2.13.3)

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

$$L_v(\lambda) = \lambda(v)$$

(2.13.5)

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

$$v \mapsto L_v$$

(2.13.6)

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

$$\Lambda(v) = (\lambda_1(v), \ldots, \lambda_n(v))$$

(2.13.7)

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

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

(2.14.1)

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

$$\{b_w : w \in V\}$$

(2.14.2)

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 $L(V)$ 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 $L(V)$ 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 $L(V)$ onto itself, because $L(V)$ 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

\[ b((T_2 \circ T_1)(v), w) = b(T_2(T_1(v)), w) = b(T_1(v), T_2^*(w)) = b(v, T_1^*(T_2^*(w))) = b(v, (T_1^* \circ T_2^*)(w)). \]

This implies that

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

so that $T \mapsto T^*$ is an opposite algebra automorphism on $L(V)$, 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

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

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

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

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

\[ (T^*)^* = T. \quad (2.14.8) \]

It follows that $T \mapsto T^*$ defines an involution on $L(V)$, 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 $b(v, w) = \overline{b(w, v)}$ (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 $b(v, w) = -\overline{b(w, v)}$ (2.15.2) is the same as saying that $i b(v, w)$ is a Hermitian form on $V$. If $b$ is any sesquilinear form on $V$, then $b(w, v)$ is a sesquilinear form on $V$ too,

$$b(v, w) + \overline{b(w, v)}$$ (2.15.3)

is a Hermitian form on $V$, and

$$b(v, w) - \overline{b(w, v)}$$ (2.15.4)

is Hermitian-antisymmetric on $V$. This permits us to express $b(v, w)$ in a unique way as $b_1(v, w) + i b_2(v, w)$, 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

$$b(T(v), w) = b(v, T(w))$$ (2.15.5)

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

$$b(T(v), w) = -b(v, T(w))$$ (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 $i T$ 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
2.15. SESQUILINEAR FORMS

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

$$b(T(v), w) = b(T^*(v), w) = b(w, (T^*)^*(v)) = b((T^*)^*(v), w).$$

(2.15.11)

Thus

$$b(T(v), w) = b(T^*(w), v) = b(w, (T^*)^*(v)) = b((T^*)^*(v), w).$$

(2.15.12)

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

\[ ab = ba = e. \]  

(3.1.1)

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

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

(3.1.2)

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

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

(3.1.3)

where $a^j$ is interpreted as being equal to $e$ when $j = 0$. In particular, if $a^{n+1} = 0,$
then it follows that \( e - a \) is invertible in \( A \), with

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

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 \( \mathcal{L}(V) \) 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 \( \mathcal{L}(V) \). 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 \( \mathcal{L}(V) \), 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 \( \mathcal{L}(V) \), 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 \( T \in \mathcal{L}(V) \) 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 \( \mathcal{L}(V) \) mentioned earlier to \( SL_n(k) \) defines a group isomorphism from \( SL_n(k) \) onto \( SL(V) \).
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

\[(3.2.1) \quad N_A(ab) \leq C N_A(a) N_A(b)\]

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

\[(3.2.2) \quad N_A(ab) = N_A(a) N_A(b)\]

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

\[(3.2.3) \quad N_A(a) \leq C N_A(a) N_A(e)\]

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

\[(3.2.4) \quad 1 \leq C N_A(e).\]

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 $\| \cdot \|_{op} = \| \cdot \|_{op, V V}$ 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

\[(3.2.5) \quad \|I\|_{op} = 1\]

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

\[(3.2.6) \quad M_a(x) = ax\]

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

\[(3.2.7) \quad N_A(M_a(x)) \leq C N_A(a) N_A(x)\]

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

\[(3.2.8) \quad \|M_a\|_{op} \leq C N_A(a)\]
for every $a \in A$, where $\| \cdot \|_{op} = \| \cdot \|_{op,A}$ is the operator seminorm on the space $\mathcal{B}\mathcal{L}(A)$ of bounded linear mappings from $A$ into itself with respect to $N_A$. If $A$ has a multiplicative identity element $e$, then

$$N_A(a) = N_A(M_a(e)) \leq \|M_a\|_{op} N_A(e) \tag{3.2.9}$$

for every $a \in A$.

Similarly, if $a \in A$, then

$$\tilde{M}_a(x) = x a \tag{3.2.10}$$

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

$$N_A(\tilde{M}_a(x)) \leq C N_A(x) N_A(a) \tag{3.2.11}$$

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

$$\|\tilde{M}_a\|_{op} \leq C N_A(a) \tag{3.2.12}$$

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

$$N_A(a) = N_A(\tilde{M}_a(e)) \leq \|\tilde{M}_a\|_{op} N_A(e) \tag{3.2.13}$$

for every $a \in A$.

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

$$\hat{N}_A(a) = \alpha N_A(a) \tag{3.2.14}$$

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

$$\hat{N}_A(a b) \leq (C/\alpha) \hat{N}_A(a) \hat{N}_A(b) \tag{3.2.15}$$

for every $a, b \in A$. In particular, this means that $\hat{N}_A$ 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 $\mathcal{B}\mathcal{L}(A)$, as in Section 2.2. This implies that $\|M_a\|_{op}$ is submultiplicative as a seminorm on $A$, because $\| \cdot \|_{op}$ is submultiplicative on $\mathcal{B}\mathcal{L}(A)$. 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

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

(3.3.1)

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

(3.3.2)

and

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

(3.3.3)

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

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

(3.3.4)

and

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

(3.3.5)

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

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

(3.3.6)

and

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

(3.3.7)

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

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

(3.3.8)

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

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

(3.3.9)

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

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

(3.3.10)
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
(3.3.12) $N_{1,\infty}(c) \leq C N_{1,\infty}(a) N_{1,\infty}(b)$.

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

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
(3.3.15) $N_{A}(c_{j,r}) \leq \max_{1 \leq l \leq n} N_{A}(a_{j,l} b_{l,r}) \leq C \max_{1 \leq l \leq n} (N_{A}(a_{j,l}) N_{A}(b_{l,r}))$

for every $j, r = 1, \ldots, n$, using (3.2.1) in the second step. It follows that
(3.3.16) $N_{\infty}(c) \leq C N_{\infty}(a) N_{\infty}(b)$.

### 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
(3.4.1) $N_{A}(e) \leq C N_{A}(a) N_{A}(a^{-1})$. 
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
\[
N_A(x^{-1} - y^{-1}) \leq C^2 N_A(x^{-1}) N_A(y^{-1}) N_A(x - y). \tag{3.4.3}
\]

Hence
\[
N_A(y^{-1}) \leq N_A(x^{-1}) + N_A(x^{-1} - y^{-1}) \leq N_A(x^{-1}) + C^2 N_A(x^{-1}) N_A(y^{-1}) N_A(x - y), \tag{3.4.4}
\]
so that
\[
(1 - C^2 N_A(x^{-1}) N_A(x - y)) N_A(y^{-1}) \leq N_A(x^{-1}). \tag{3.4.5}
\]

If
\[
C^2 N_A(x^{-1}) N_A(x - y) < 1, \tag{3.4.6}
\]
then we get that
\[
N_A(y^{-1}) \leq (1 - C^2 N_A(x^{-1}) N_A(x - y))^{-1} N_A(x^{-1}). \tag{3.4.7}
\]

Combining this with (3.4.3), we obtain that
\[
N_A(x^{-1} - y^{-1}) \leq C^2 (1 - C^2 N_A(x^{-1}))^{-1} N_A(x^{-1})^2 N_A(x - y) \tag{3.4.8}
\]
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
\[
N_A(x^{-1}) = N_A(y^{-1}) \tag{3.4.9}
\]
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
\[
C^2 N_A(x^{-1}) N_A(y^{-1}) N_A(x - y) < N_A(y^{-1}). \tag{3.4.10}
\]

Combining this with (3.4.3), we get that
\[
N_A(x^{-1} - y^{-1}) < N_A(y^{-1}). \tag{3.4.11}
\]

This implies (3.4.9) in this situation, as in (1.8.7). It follows that
\[
N_A(x^{-1} - y^{-1}) \leq C^2 N_A(x^{-1})^2 N_A(x - y) \tag{3.4.12}
\]
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
\[
e^{-1} - y^{-1} = (y - e) y^{-1}, \tag{3.4.13}
\]
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
\[ (1 - C N_A(y - e)) N_A(y^{-1}) \leq N_A(e). \] (3.4.16)

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
3.5. BANACH ALGEBRAS

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

\[
\| e \| = 1
\]

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

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

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

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

\[
\| a^j \| \leq C^{j-1} \| a \|^j
\]

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

\[
C \| a \| < 1,
\]

so that

\[
\lim_{j \to \infty} \| a^j \| \to 0,
\]

by (3.5.4). We also get that

\[
\sum_{j=0}^{\infty} \| a^j \|
\]

converges as an infinite series of nonnegative real numbers, because

\[
\sum_{j=0}^{\infty} C^j \| a \|^j
\]

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

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

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

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

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

\[
y = x - (x - y) = x (e - x^{-1} (x - y)).
\]
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 \( \|a^{j_0 + r}\| \) in terms of \( C^{j_0 - 1} \|a^r\| \) 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
\[ N_V(v) \leq C N_W(T_1(v)) + C \|T_1 - T\|_{op,V_W} N_V(v) \]
for every \( v \in V \), where the operator seminorm \( \| \cdot \|_{op,V_W} \) is as defined in Section 1.9. Thus
\[ (1 - C \|T_1 - T\|_{op,V_W}) N_V(v) \leq C N_W(T_1(v)) \]
3.7. ISOMETRIC LINEAR MAPPINGS

for every \( v \in V \). If

\[ C \| T_1 - T \|_{op, V W} < 1, \]

then it follows that

\[ N_V(v) \leq C \left( 1 - C \| T_1 - T \|_{op, V W} \right)^{-1} N_W(T_1(v)) \]

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

\[ N_V(v) \leq C \max(N_W(T_1(v)), N_W(T_1(v) - T(v))) \]

for every \( v \in V \), by (3.6.2). Suppose for the moment that

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

for every \( v \in V \) with \( N_V(v) > 0 \). It follows that

\[ N_V(v) \leq C N_W(T_1(v)) \]

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

\[ N_W(T_1(v) - T(v)) \leq \| T_1 - T \|_{op, V W} N_V(v) \]

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 \( BL(V) \), 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

\[ N_W(T(v)) = N_V(v) \]
for every \( v \in V \). Of course, this is the same as saying that
\[
N_W(T(v)) \leq N_V(v)
\]
and
\[
N_V(v) \leq N_W(T(v))
\]
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
\[
\|T\|_{op,V,W} \leq 1,
\]
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
\[
N_Z((T_2 \circ T_1)(v)) = N_Z(T_2(T_1(v))) = N_W(T_1(v)) = N_V(v)
\]
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
\[
\|T^{-1}\|_{op,W,V} \leq 1,
\]
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
\[
N_W(T_1(v) - T(v)) < N_V(v)
\]
for every \( v \in V \) with \( N_V(v) > 0 \). Under these conditions, we have that
\[
N_V(v) \leq N_W(T_1(v))
\]
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
\[
\|T_1 - T\|_{op,V,W} \leq 1,
\]
then
\[ \|T_1\|_{op,WV} \leq 1, \]
because of (3.7.4) and the hypothesis that \( N_W \) be a semi-ultranorm on \( W \). If
\[ \|T_1 - T\|_{op,WV} < 1, \]
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
\[ \langle T(u), T(v) \rangle_W = \langle u, v \rangle_V \]
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
\[ \langle T(u), T(v) \rangle_W = \langle u, T^*(T(v)) \rangle_V \]
for every \( u, v \in V \). Thus \( T \) is an isometric linear mapping from \( V \) into \( W \) if and only if
\[ T^* \circ T = I_V, \]
where \( I_V \) is the identity mapping on \( V \). This is the same as saying that
\[ T^* = T^{-1} \]
when \( T \) maps \( V \) onto \( W \).

If \( T \) is any bounded linear mapping from \( V \) into \( W \), then it is well known that
\[ \|T^* \circ T\|_{op,VV} = \|T\|_{op,WV}^2, \]
where these operator norms are taken with respect to \( \| \cdot \|_V \) and \( \| \cdot \|_W \), as appropriate. More precisely,
\[ \|T^* \circ T\|_{op,VV} \leq \|T\|_{op,VW} \|T^*\|_{op,WV} = \|T\|_{op,WV}^2, \]
using (1.11.15) in the second step. We also have that
\[
\|T(v)\|^2_W = \langle v, T^*(T(v)) \rangle_V \leq \|v\|_V \|T^* \circ T\|_{op,VV} \|v\|^2_V
\]
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
\[
\|T\|_{op,VW}^2 \leq \|T^* \circ T\|_{op,VV},
\]
(3.8.8)
as desired.

Observe that \(w \in W\) satisfies
\[
T^*(w) = 0 \text{ if and only if } \langle T(v), w \rangle_W = 0
\]
(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
\[
\|v\|_V \leq C \|T(v)\|_W
\]
(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
\[
\beta_2(\phi(a_1), \phi(b_1)) = \beta_1(a_1, b_1)
\]
(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
\[
\beta_1(\phi^{-1}(a_2), \phi^{-1}(b_2)) = \beta_2(a_2, b_2)
\]
(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

$$\beta_3(\phi_2(\phi_1(a_1)), \phi_2(\phi_1(b_1))) = \beta_2(\phi_1(a_1), \phi_1(b_1)) = \beta_1(a_1, b_1)$$

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

$$\beta_2(\phi(a), \phi(a)) = \beta_1(a, a)$$

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

$$\beta_1(a + b, a + b) = \beta_1(a, a) + \beta_1(a, b) + \beta_1(b, a) + \beta_1(b, b)$$

and

$$\beta_2(\phi(a + b), \phi(a + b)) = \beta_2(\phi(a) + \phi(b), \phi(a) + \phi(b)) = \beta_2(\phi(a), \phi(a)) + \beta_2(\phi(a), \phi(b)) + \beta_2(\phi(b), \phi(a)) + \beta_2(\phi(b), \phi(b)).$$

It follows that

$$\beta_2(\phi(a), \phi(b)) + \beta_2(\phi(b), \phi(a)) = \beta_1(a, b) + \beta_1(b, a)$$

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

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

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 $A$. 

---

3.9. PRESERVING BILINEAR FORMS

63
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

$$b_2(T(v_1), w_2) = b_1(v_1, T^*(w_2)) \tag{3.10.1}$$

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 $V_2$ also has finite dimension, and that $b_3$ is nondegenerate on $V_2$. 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 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

$$b_3((T_2 \circ T_1)(v_1)), w_3) = b_3(T_2(T_1(v_1)), w_3) = b_2(T_1(v_1), T_2^*(w_3)) \tag{3.10.2}$$

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

$$b_2(T(v_1), w_2) = b_1(T^*(w_2), v_1) = b_2(w_2, (T^*)^*(v_1)) = b_2((T^*)^*(v_1), w_2) \tag{3.10.3}$$

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

$$b_2(T(v_1), w_2) = b_1(T^*(w_2), v_1) = b_2(w_2, (T^*)^*(v_1)) = b_2((T^*)^*(v_1), w_2) \tag{3.10.4}$$

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

$$b_2(T(v_1), w_2) = -b_1(T^*(w_2), v_1) = -b_2(w_2, (T^*)^*(v_1)) = b_2((T^*)^*(v_1), w_2) \tag{3.10.5}$$
for every $v_1 \in V_1$ and $w_2 \in V_2$. In both cases, it follows that

$$\text{(3.10.6)} \quad (T^*)^* = T.$$  

If one of $b_1$ and $b_2$ is symmetric, and the other is antisymmetric, then

$$\text{(3.10.7)} \quad (T^*)^* = -T,$$

by the analogous argument.

If $T$ is any linear mapping from $V_1$ into $V_2$, then

$$\text{(3.10.8)} \quad b_2(T(v_1), T(w_1)) = b_1(v_1, T^*(T(w_1)))$$

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

$$\text{(3.10.9)} \quad T^* \circ T = I_{V_1},$$

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

$$\text{(3.10.10)} \quad T^* = T^{-1}.$$  

### 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

$$\text{(3.11.1)} \quad b_2(T(v_1), T(w_1)) = b_1(v_1, w_1)$$

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

$$\text{(3.11.2)} \quad b_1(T^{-1}(v_2), T^{-1}(w_2)) = b_2(v_2, w_2)$$

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

$$\text{(3.11.3)} \quad b_2(T(u_1), T(u_1)) = b_1(u_1, u_1)$$

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 $u_1 = v_1 + iw_1$. Of course, (3.11.1) implies (3.11.3), by taking $v_1, w_1 = u_1$. 


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
\begin{equation}
(3.11.4) \quad b(T(v), T(w)) = b(v, w)
\end{equation}
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
\begin{equation}
(3.11.5) \quad b_2(T(v_1), w_2) = b_1(v_1, T^*(w_2))
\end{equation}
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 $V_2$. 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
\begin{equation}
(3.11.6) \quad (T_2 \circ T_1)^* = T_1^* \circ T_2^*,
\end{equation}
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 $V_2$ 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
\begin{equation}
(3.11.7) \quad b_2(T(v_1), w_2) = \overline{b_1(T^*(w_2), v_1)} = \overline{b_2(w_2, (T^*)^*(v_1))} = b_2((T^*)^*(v_1), w_2)
\end{equation}
for every $v_1 \in V_1$ and $w_2 \in V_2$. This implies that
\begin{equation}
(3.11.8) \quad (T^*)^* = T
\end{equation}
in this situation.
3.12. BILINEAR FORMS AND MATRICES

As before,
\begin{equation}
(3.11.9) \quad b_2(T(v_1), T(w_1)) = b_1(v_1, T^*(T(w_1)))
\end{equation}
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
\begin{equation}
(3.11.10) \quad T^* \circ T = I_{V_1},
\end{equation}
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
\begin{equation}
(3.11.11) \quad T^* = T^{-1}.
\end{equation}
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
\begin{equation}
(3.12.1) \quad \beta(x, y) = \sum_{j=1}^{n} \sum_{l=1}^{n} \beta_{j,l} x_l y_j
\end{equation}
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
\begin{equation}
(3.12.2) \quad \beta_{l,j} = \beta_{j,l}
\end{equation}
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
\begin{equation}
(3.12.3) \quad \beta_{l,j} = -\beta_{j,l}
\end{equation}
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
\begin{equation}
(3.12.4) \quad \beta(x, x) = 0
\end{equation}
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
\begin{equation}
(3.12.5) \quad \beta_{j,j} = 0
\end{equation}
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

\[
(T_a(x))_j = \sum_{l=1}^{n} a_{j,l} \, x_l \tag{3.12.6}
\]

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

\[
\beta(T_a(x), y) = \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{r=1}^{n} \beta_{r,j} a_{j,l} \, x_l \, y_r \tag{3.12.7}
\]

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,

\[
\beta(x, T_a(y)) = \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{r=1}^{n} \beta_{j,l} a_{j,r} \, x_l \, y_r \tag{3.12.8}
\]

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

\[
\sum_{j=1}^{n} \beta_{r,j} a_{j,l} = \sum_{j=1}^{n} \beta_{j,l} a_{j,r} \tag{3.12.9}
\]

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

\[
\sum_{j=1}^{n} \beta_{r,j} a_{j,l} = -\sum_{j=1}^{n} \beta_{j,l} a_{j,r} \tag{3.12.10}
\]

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

We also have that

\[
\beta(T_a(x), T_a(y)) = \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{m=1}^{n} \sum_{r=1}^{n} \beta_{m,j} a_{j,l} \, a_{m,r} \, x_l \, y_r \tag{3.12.11}
\]

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

\[
\sum_{j=1}^{n} \sum_{m=1}^{n} \beta_{m,j} a_{j,l} \, a_{m,r} = \beta_{r,l} \tag{3.12.12}
\]
3.12. BILINEAR FORMS AND MATRICES

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

$$\beta(z, w) = \sum_{j=1}^{n} \sum_{l=1}^{n} \beta_{j,l} \overline{z_l} w_j \quad (3.13.1)$$

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

$$\beta_{l,j} = \overline{\beta_{j,l}} \quad (3.13.2)$$

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

$$(T_a(z))_j = \sum_{l=1}^{n} a_{j,l} z_l \quad (3.13.3)$$

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 $\mathcal{L}(\mathbb{C}^n)$ 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

$$(a^*)_{j,l} = \overline{a_{l,j}} \quad (3.13.4)$$

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

$$\beta(T_a(z), w) = \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{r=1}^{n} \beta_{r,j} a_{j,l} z_l \overline{w_r} \quad (3.13.5)$$

for every $z, w \in \mathbb{C}^n$. Similarly,

$$\beta(z, T_a(w)) = \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{r=1}^{n} \beta_{j,l} \overline{a_{j,r}} z_l \overline{w_r} \quad (3.13.6)$$

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

$$\sum_{j=1}^{n} \beta_{r,j} a_{j,l} = \sum_{j=1}^{n} \beta_{j,l} \overline{a_{j,r}} \quad (3.13.7)$$
for every \( l, r = 1, \ldots, n \). If we also use \( \beta \) to denote \((\beta_{j,l})\) as an element of \( M_n(\mathbb{C}) \), 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 \( z, w \in \mathbb{C}^n \). It follows that \( T_a \) preserves \( \beta(\cdot, \cdot) \), as in Section 3.11, if and only if

\[
(3.13.10) \quad \sum_{j=1}^{n} \sum_{m=1}^{n} \beta_{m,j} a_{j,l} \overline{a_{m,r}} = \beta_{r,l}
\]

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 \( M_n(\mathbb{C}) \).

If \( b \in M_n(\mathbb{C}) \), 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 \( z, w \in \mathbb{C}^n \), as before. Hence

\[
(3.13.13) \quad \beta(T_a(z), w) = \beta(z, T_b(w))
\]

for every \( z, w \in \mathbb{C}^n \) 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 \( M_n(\mathbb{C}) \). Suppose that \((\beta_{j,l})\) is invertible as an element of \( M_n(\mathbb{C}) \), which is equivalent to the nondegeneracy of \( \beta(\cdot, \cdot) \) as a sesquilinear form on \( \mathbb{C}^n \), 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 \textit{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 \textit{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

\[(3.14.11) \quad (xy)^* \beta xy = y^* x^* \beta xy = y^* \beta y = \beta.\]

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

\[(\phi_\beta(\phi_\beta(x))) = \beta^{-1} x (\beta^*)^{-1} \beta\]

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 \( i x \) 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

\[(3.15.1) N_A(e) = 1.\]

If \( x \) is an invertible element of \( A \), then it follows that

\[(3.15.2) N_A(x) N_A(x^{-1}) \geq 1.\]

In particular, if we have that

\[(3.15.3) N_A(x), N_A(x^{-1}) \leq 1,\]

then

\[(3.15.4) N_A(x) = N_A(x^{-1}) = 1.\]

Let \( y \) be another invertible element of \( A \) such that

\[(3.15.5) N_A(y), N_A(y^{-1}) \leq 1.\]

Thus \( xy \) is an invertible element of \( A \) too,

\[(3.15.6) N_A(x, y) \leq N_A(x) N_A(y) \leq 1,\]

and

\[(3.15.7) N_A((xy)^{-1}) = N_A(y^{-1} x^{-1}) \leq N_A(x^{-1}) N_A(y^{-1}) \leq 1.\]
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

$$N_A(x - y) < 1.$$  \hspace{1cm} (3.15.8)

This implies that

$$N_A(y) \leq \max(N_A(x), N_A(x - y)) \leq 1.$$  \hspace{1cm} (3.15.9)

We also have that

$$N_A(y^{-1}) = N_A(x^{-1}) \leq 1,$$  \hspace{1cm} (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

$$N_A(x^*) \leq C_1 N_A(x)$$  \hspace{1cm} (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

$$N_A(x) \leq C_2 N_A(x^*)$$  \hspace{1cm} (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

$$N_A(x^*) = N_A(x)$$  \hspace{1cm} (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

\[(4.1.1) \quad \tilde{N}_l(v) = N_l(v_l)\]

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 $\prod_{j \in I} V_j$. 

76
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

\[ \tilde{N}^1(v) = \sum_{j \in I} N_j(v_j) \tag{4.1.2} \]

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

\[ \tilde{N}^\infty(v) = \max_{j \in I} N_j(v_j) \tag{4.1.3} \]

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

\[ \tilde{N}^\infty(v) \leq \tilde{N}^1(v) \tag{4.1.4} \]

for every $v \in \bigoplus_{j \in I} V_j$. If $I$ has only finitely many elements, then

\[ \tilde{N}^1(v) \leq (\#I) \tilde{N}^\infty(v) \tag{4.1.5} \]

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,

\[ \tilde{N}^1_a(v) = \sum_{j \in I} a_j N_j(v_j) \tag{4.1.6} \]

defines a seminorm on $\bigoplus_{j \in I} V_j$ with respect to $\| \cdot \|$ on $k$. Similarly,

\[ \tilde{N}^\infty_a(v) = \max_{j \in I} (a_j N_j(v_j)) \tag{4.1.7} \]

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

\[ \tilde{N}^\infty_a(v) \leq \tilde{N}^1_a(v) \tag{4.1.8} \]

for every $v \in \bigoplus_{j \in I} V_j$. If $I$ has only finitely many elements, then

\[ \tilde{N}^1_a(v) \leq \left( \sum_{j \in I} a_j \right) \tilde{N}^\infty(v) \tag{4.1.9} \]

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

$$c_n = \sum_{j=0}^{n} \beta(a_j, b_{n-j})$$

for each nonnegative integer $n$. It is easy to see that

$$\sum_{n=0}^{\infty} c_n = \beta\left(\sum_{j=0}^{\infty} a_j, \sum_{l=0}^{\infty} b_l\right),$$

at least formally. More precisely, suppose for the moment that there are nonnegative 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 $\sum_{j=0}^{J} a_j$, $\sum_{l=0}^{L} b_l$, 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 $a \cdot b$. The Cauchy product of these series is defined to be the series $\sum_{n=0}^{\infty} c_n$, where

$$c_n = \sum_{j=0}^{n} a_j b_{n-j},$$

as in (4.2.1). Thus

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{j=0}^{\infty} a_j\right) \left(\sum_{l=0}^{\infty} b_l\right),$$

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

$$\left(\sum_{j=0}^{J} a_j\right) \left(\sum_{l=0}^{L} b_l\right) \leq \sum_{n=0}^{J+L} c_n.$$

Similarly, if $N$ is a nonnegative integer, then

$$\sum_{n=0}^{N} c_n \leq \left(\sum_{j=0}^{N} a_j\right) \left(\sum_{l=0}^{N} b_l\right).$$
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 \(|·|\), and let \( A, B, \) and \( C \) be vector spaces over \( k \) with norms \( N_A, N_B, \) and \( N_C, \) respectively, with respect to \(|·|\) 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

\[
\sum_{j=0}^{\infty} a_j \quad \text{and} \quad \sum_{l=0}^{\infty} b_l
\]

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 \( f(T) \) in \( T \) with coefficients in \( A \) can be expressed as

\[
f(T) = \sum_{j=0}^{\infty} f_j T^j,
\]

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

\[
f(T) = \sum_{j=0}^{n} f_j T^j
\]

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

\[
f(T) \mapsto f_0
\]

is a module homomorphism from \( A[[T]] \) onto \( A \).

If \( f(T) \in A[[T]] \) and \( l \) is a nonnegative integer, then

\[
f(T) T^l = \sum_{j=0}^{\infty} f_j T^{j+l} = \sum_{j=l}^{\infty} f_{j-l} T^j
\]

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

\[
f(T) \mapsto f(T) T^l
\]
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 if for each nonnegative integer $j$ there is a positive integer $L_j$ such that

\begin{equation}
\hat{f}_{l,j} = f_j
\end{equation}

for every $l \geq L_j$. This implies that $f_l(T) - f(T)$ vanishes to order $n \geq 0$ 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 $\hat{f}_{l,j}$ 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

\begin{equation}
\sum_{l=1}^{n} a_l(T)
\end{equation}
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

\[(4.4.3) \quad \sum_{l=1}^{\infty} a_l(T)\]

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 $\{\alpha a_l(T)\}_{l=1}^{\infty}$ eventually agrees termwise with 0 too, and

\[(4.4.4) \quad \sum_{l=1}^{\infty} \alpha a_l(T) = \alpha \sum_{l=1}^{\infty} a_l(T).\]

If $r$ is a nonnegative integer, then $\{a_l(T)T^r\}_{l=1}^{\infty}$ eventually agrees termwise with 0 as well, and

\[(4.4.5) \quad \sum_{l=1}^{\infty} a_l(T)T^r = \left(\sum_{l=1}^{\infty} a_l(T)\right)T^r.\]

If $\{b_l(T)\}_{l=1}^{\infty}$ is another sequence of elements of $A[[T]]$ that eventually agrees termwise with 0, then $\{a_l(T) + b_l(T)\}_{l=1}^{\infty}$ eventually agrees termwise with 0, and

\[(4.4.6) \quad \sum_{l=1}^{\infty} (a_l(T) + b_l(T)) = \sum_{l=1}^{\infty} a_l(T) + \sum_{l=1}^{\infty} b_l(T).\]

Of course, one can deal with sequences and series that start with $l = 0$ in the same way.

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 $f(T) = \sum_{l=0}^{\infty} f_l T^l \in A[[T]]$, then $\phi(f_l) \in B[[T]]$ for each $l \geq 0$, and $\phi(f_l) T^l$ automatically vanishes to order $l - 1$ for every $l \geq 1$. In particular, $\{\phi(f_l) T^l\}_{l=0}^{\infty}$ eventually agrees termwise with 0, and we put

\[(4.4.7) \quad \phi(f(T)) = \sum_{l=0}^{\infty} \phi(f_l) T^l,\]

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

\[(4.4.8) \quad \phi(f(T) T^r) = \phi(f(T)) T^r\]
for every \( f(T) \in A[[T]] \) and nonnegative integer \( r \). If \( f(T) = \sum_{l=0}^{\infty} f_l T^l \in A[[T]] \) and \( j \) is a nonnegative integer, then the total coefficient of \( T^j \) in (4.5.7) is the sum of the coefficients of \( T^{j-l} \) 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
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 \( B[[T]] \) that eventually agree termwise with \( f(T) \) and \( g(T) \), 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 0. Thus \( \sum_{m=0}^\infty a_m(T) \) and \( \sum_{r=0}^\infty b_r(T) \) can be defined as elements of \( A[[T]] \) and \( B[[T]] \), respectively, as in the previous section. Using the extension of \( \beta \) to \( A[[T]] \times B[[T]] \) defined earlier, we get that \( \beta(a_m(T), b_r(T)) \) is defined as an element of \( C[[T]] \) for all \( m, r \geq 0 \). Put

\[
(4.5.5) \quad c_N(T) = \sum_{m=0}^N \beta(a_m(T), b_{N-m}(T))
\]

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

\[
(4.5.6) \quad \sum_{N=0}^\infty c_N(T) = \beta\left( \sum_{m=0}^\infty a_m(T), \sum_{r=0}^\infty b_r(T) \right)
\]

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

\[
(4.5.7) \quad \sum_{l=0}^n \beta(f_{n-l}, g_l).
\]

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

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

for every \( a \in A \), and let us check that the extension of \( \beta \) to \( A[[T]] \times A[[T]] \) satisfies

\[
(4.5.9) \quad \beta(f(T), f(T)) = 0
\]

for every \( f(T) = \sum_{j=0}^\infty f_j T^j \in A[[T]] \). To do this, it suffices to verify that

\[
(4.5.10) \quad \sum_{j=0}^n \beta(f_j, f_{n-j}) = 0
\]
for every nonnegative integer \( n \). Remember that (4.5.8) implies that \( \beta \) is antisymmetric 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 \( a \cdot b \). Let \( f(T) = \sum_{j=0}^{\infty} f_j T^j \) and \( g(T) = \sum_{i=0}^{\infty} g_i T^i \) 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 subalgebra 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 \( k[T] \).
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 from $B[T]$, 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 $(A, [\cdot, \cdot]_A)$ 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

\begin{equation}
(4.6.3) \quad h_n = \sum_{j=0}^{n} [f_j, g_{n-j}]_A
\end{equation}

for each $n \geq 0$, and we put

\begin{equation}
(4.6.4) \quad [f(T), g(T)]_{A[[T]]} = h(T) = \sum_{n=0}^{\infty} h_n T^n.
\end{equation}

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 $[\cdot, \cdot]_{A[[T]]}$ from the Jacobi identity for $[\cdot, \cdot]_A$. 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

\begin{equation}
(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}
\end{equation}
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 (4.7.2) to be the same as in (4.7.1) 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

\[
(e - a(T)) \sum_{l=0}^{\infty} a(T)^l = \left( \sum_{l=0}^{\infty} a(T)^l \right) (e - a(T)) = e.
\]

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 \( f(T) = \sum_{j=0}^{\infty} f_j T^j \) 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

\[
\phi(f(T) T) = \phi(f(T)) T
\]

for every \( f(T) \in A[[T]] \). This implies that

\[
\phi(f(T) T^r) = \phi(f(T)) T^r
\]
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 \( k \) 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 \( A \) 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 \( \text{Hom}_k(A,T) \) 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 \( k \) 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

\[
\text{Hom}_{k[[T]]}(A[[T]],B[[T]]),
\]
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

\[
(4.8.6) \quad \beta(f(T), T, g(T)) = \beta((f(T), g(T)) T)
\]

for every $f(T) \in A[T]$ and $g(T) \in B[T]$, then

\[
(4.8.7) \quad \beta(f(T), T^r, g(T), T^m) = \beta(f(T), g(T)) T^{r+m}
\]

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

\[
(4.8.8) \quad \beta\left(\sum_{j=0}^{r} f_j T^j, \sum_{i=0}^{r} g_i T^i\right).
\]

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
CHAPTER 4. FORMAL POWER SERIES

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

\[
(4.9.1) \quad \phi(T) = \sum_{j=0}^{\infty} \phi_j T^j
\]

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

\[
(4.9.2) \quad (\phi(T))(a) = \sum_{j=0}^{\infty} \phi_j(a) T^j
\]

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

\[
(4.9.3) \quad \phi(T) = \sum_{j=0}^{n} \phi_j T^j
\]

is a formal polynomial in \(T\) with coefficients in \(\text{Hom}_k(A, B)\), then

\[
(4.9.4) \quad (\phi(T))(a) = \sum_{j=0}^{n} \phi_j(a) T^j
\]

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

\[(\phi(T))(a_m) = \sum_{j=0}^\infty \phi_j(a_m) T^j\]  

(4.10.1)

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

\[(\phi(T))(a(T)) = \sum_{m=0}^\infty (\phi(T))(a_m) T^m\]  

(4.10.2)

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

\[E(\phi, a) = \phi(a)\]  

(4.10.3)

for every $\phi \in \text{Hom}_k(A, B)$ and $a \in A$, which is the natural evaluation mapping

\[E: \text{Hom}_k(A, B) \times A \rightarrow B.\]  

(4.10.4)

This mapping is bilinear over $k$, and can be extended to a mapping

\[E_n: (\text{Hom}_k(A, B))[[T]] \times A[[T]] \rightarrow B[[T]],\]  

(4.10.5)

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

\[E_n(\phi(T), a(T)) = \sum_{j=0}^n \phi_j(a_{n-j})\]  

(4.10.6)
for each nonnegative integer $n$. This is an element of $B$ for every $n \geq 0$, so that

$$E(\phi(T), a(T)) = \sum_{n=0}^{\infty} E_n(\phi(T), a(T)) T^n$$

(4.10.7)

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

$$\psi(T) = \sum_{l=0}^{\infty} \psi_l T^l$$

(4.10.8)

from $(\text{Hom}_k(A, B))[T] \times A[T]$ 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

$$\psi(T) = \sum_{l=0}^{\infty} \psi_l T^l$$

(4.10.9)

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

$$(\psi(T) \circ \phi(T))_n = \sum_{l=0}^{n} \psi_l \circ \phi_{n-l}$$

(4.10.10)

is an element of $\text{Hom}_k(A, C)$ for every nonnegative integer $n$. Put

$$\psi(T) \circ \phi(T) = \sum_{n=0}^{\infty} (\psi(T) \circ \phi(T))_n T^n,$$

(4.10.11)

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

$$\psi(T) \circ \phi(T)$$

(4.10.12)

from $\text{Hom}_k(A, B) \times \text{Hom}_k(B, C)$ 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

$$\psi(T) \circ \phi(T)$$

(4.10.13)

from $(\text{Hom}_k(A, B))[T] \times (\text{Hom}_k(B, C))[T]$ into $\text{Hom}_k(A, C))[T]$

and

$$\psi(T) \circ \phi(T)$$

(4.10.14)

from $(\text{Hom}_k(A, B))[T] \times (\text{Hom}_k(B, C))[T]$ into $\text{Hom}_k(A, C))[T]$,
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 \( A[[T]] \) 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

\[
\rho_a(b) = \beta(a, b)
\]

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

\[
a \mapsto \rho_a
\]

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 \( \text{Hom}_k[[T]](B[[T]], C[[T]]) \) and \( (\text{Hom}_k(B, C))[T] \), as modules over \( k[[T]] \), as in Sections 4.8 and 4.9. The isomorphism with \( \text{Hom}_k[[T]](B[[T]], C[[T]]) \) 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 \( A[[T]] \).
Now let $\beta$ be a mapping from $A \times B$ into $C[T]$ that is bilinear over $k$. If $a \in A$, then (4.11.1) defines $\rho_a$ as a mapping from $B$ into $C[T]$ that is linear over $k$, and hence an element of $\text{Hom}_k(B, C[T])$. Similarly, (4.11.2) defines a mapping from $A$ into $\text{Hom}_k(B, C[T])$ that is linear over $k$. There is a natural isomorphism between $\text{Hom}_k(B, C[T])$ and $\text{Hom}_{k[T]}(B[T], C[T])$, as modules over $k[T]$, as in Section 4.8. This permits one to identify $\rho_a$ with a homomorphism from $B[T]$ into $C[T]$, as modules over $k[T]$.

As in Section 4.9, there is a natural injection from $(\text{Hom}_k(B, C))[T]$ into $\text{Hom}_k(B, C[T])$ that is linear over $k[T]$. If (4.11.2) corresponds to a mapping from $A$ into $(\text{Hom}_k(B, C))[T]$ that is linear over $k$, then this mapping can be extended to one from $A[T]$ into $(\text{Hom}_k(B, C))[T]$ that is linear over $k[T]$ in the usual way. Otherwise, (4.11.2) can still be extended to a mapping from $A[T]$ into $\text{Hom}_k(B, C[T])$ or equivalently $\text{Hom}_{k[T]}(B[T], C[T])$ that is linear over $k[T]$. This basically just uses the fact that $\text{Hom}_k(B, C[T])$ or equivalently $\text{Hom}_{k[T]}(B[T], C[T])$ is a module over $k[T]$. Of course, one can also look at this in terms of extending (4.11.1) in $a$ for $b$ fixed, as mentioned at the beginning of the section.

### 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. In particular, $A[[T]]$ 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

$$
\phi(a) \phi(a') = \left( \sum_{j=0}^{\infty} \phi_j(a) T^j \right) \left( \sum_{l=0}^{\infty} \phi_l(a') T^l \right) = \sum_{n=0}^{\infty} \left( \sum_{j=0}^{n} \phi_j(a) \phi_{n-j}(a') \right) T^n.
$$
4.12. EXTENDING ALGEBRA HOMOMORPHISMS

It follows that (4.12.2) holds if and only if

\[(4.12.5) \quad \phi_n(a a') = \sum_{j=0}^{n} \phi_j(a) \phi_{n-j}(a')\]

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

\[(4.12.6) \quad \phi(f(T)) = \sum_{j=0}^{\infty} \phi(f_j) T^j\]

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

\[(4.12.7) \quad h_n = \sum_{j=0}^{n} f_j g_{n-j}\]

for every \(n \geq 0\), as in Section 4.6. Thus

\[(4.12.8) \quad \phi(f(T) g(T)) = \phi(h(T)) = \sum_{n=0}^{\infty} \phi(h_n) T^n,\]

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

\[(4.12.9) \quad \phi(h_n) = \sum_{j=0}^{n} \phi(f_j g_{n-j}) = \sum_{j=0}^{n} \phi(f_j) \phi(g_{n-j})\]

for each \(n \geq 0\). This implies that

\[(4.12.10) \quad \phi(f(T) g(T)) = \sum_{n=0}^{\infty} \left( \sum_{j=0}^{n} \phi(f_j) \phi(g_{n-j}) \right) T^n = \phi(f(T)) \phi(g(T)).\]

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.

Put $\alpha(T) = a(T) - e = \sum_{j=1}^{\infty} a_j T^j$ and $\beta(T) = b(T) - e = \sum_{j=1}^{\infty} b_j T^j$, so that $
abla(T) = e + \alpha(T)$ and $\gamma(T) = e + \beta(T)$. Thus

\begin{equation}
\tag{4.13.1}
\nabla(T) \gamma(T) = e + \alpha(T) + \beta(T) + \alpha(T) \beta(T).
\end{equation}

As in Section 4.7,

\begin{equation}
\tag{4.13.2}
\nabla(T)^{-1} = \sum_{l=0}^{\infty} (-\alpha(T))^l = e + \sum_{l=1}^{\infty} (-\alpha(T))^l
\end{equation}

and

\begin{equation}
\tag{4.13.3}
\gamma(T)^{-1} = \sum_{l=0}^{\infty} (-\beta(T))^l = e + \sum_{l=1}^{\infty} (-\beta(T))^l.
\end{equation}

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

\begin{equation}
\tag{4.13.4}
\nabla(T) \gamma(T) \nabla(T)^{-1} \gamma(T)^{-1} = e + O(T),
\end{equation}

because $\nabla(T), \gamma(T), \nabla(T)^{-1}, \gamma(T)^{-1} = e + O(T)$. More precisely,

\begin{equation}
\tag{4.13.5}
\nabla(T) \gamma(T) \nabla(T)^{-1} \gamma(T)^{-1} = e + O(T^2),
\end{equation}

because

\begin{equation}
\tag{4.13.6}
\nabla(T) = e + a_1 T + O(T^2), \quad \gamma(T) = e + b_1 T + O(T^2),
\end{equation}

and

\begin{equation}
\tag{4.13.7}
\nabla(T)^{-1} = e - a_1 T + O(T^2), \quad \gamma(T)^{-1} = e - b_1 T + O(T^2).
\end{equation}

We also have that

\begin{equation}
\tag{4.13.8}
\nabla(T) = e + a_1 T + a_2 T^2 + O(T^3), \quad \gamma(T) = e + b_1 T + b_2 T^2 + O(T^3),
\end{equation}

and

\begin{equation}
\tag{4.13.9}
\nabla(T) \gamma(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).
\end{equation}

Using (4.13.2), we get that

\begin{equation}
\tag{4.13.10}
\nabla(T)^{-1} = e - \nabla(T)^2 + O(T^3)
= e - a_1 T - a_2 T^2 + a_1^2 T^2 + O(T^3).
\end{equation}
Similarly,
(4.13.11) \[ b(T)^{-1} = e - b_1 T - b_2 T^2 + b_1^2 T^2 + O(T^3). \]

It follows that
(4.13.12) \[
a(T)^{-1} b(T)^{-1} = e - a_1 T - b_1 T - a_2 T^2 - b_2 T^2 + a_1^2 T^2 + b_1^2 T^2 + a_1 b_1 T^2 + O(T^3).
\]

Combining (4.13.9) and (4.13.12), it is not difficult to verify that
(4.13.13) \[
a(T) b(T) a(T)^{-1} b(T)^{-1} = e + a_1 b_1 T^2 - b_1 a_1 T^2 + O(T^3)
\]

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) \[ a(T)^* = \sum_{j=0}^{\infty} a_j^* T^j \]
defines an element of \( A[[T]] \), which is in \( A[T] \) when \( a(T) \in A[T] \). Of course,
(4.14.2) \[ a(T) \mapsto a(T)^* \]
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) \[ y^* y = (e + y_0^*) (e + y_0) = e + y_0 + y_0^* + y_0^* y_0. \]

Suppose that \( y_1, y_2 \in A \) satisfy
(4.14.4) \[ y_0 = y_1 + y_2, \quad y_1^* = y_1, \quad y_2^* = -y_2, \quad \text{and} \quad y_1 y_2 = y_2 y_1. \]

Under these conditions, \( y_0^* = y_1 - y_2 \), and
(4.14.5) \[ y^* y = e + 2 \cdot y_1 + (y_1 - y_2) (y_1 + y_2) = e + 2 \cdot y_1 + y_1^2 - y_2^2. \]
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

\[
(4.14.6) \quad a(T)^* a(T) = e + \alpha(T) + \alpha(T)^* \alpha(T).
\]

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 \( a(T)^* a(T) = e + O(T^2) \), in the notation of the previous section. Of course, (4.14.7) is the same as saying that

\[
(4.14.9) \quad a(T)^* = a(T)^{-1},
\]

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 \( \beta_j \) 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

\[ \epsilon_n^{n+1} = 0. \]  

(5.1.1)

The elements of $A[\epsilon_n]$ may be expressed as formal sums of the form

\[ a = a_0 + a_1 \epsilon_n + \cdots + a_n \epsilon_n^n, \]  

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

\[ a \epsilon_n = a_0 \epsilon_n + \cdots + a_{n-1} \epsilon_n^n \]  

(5.1.3)

defines an element of $A[\epsilon_n]$ as well. This defines

\[ a \mapsto a \epsilon_n \]  

(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 $\epsilon_n^0$ 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

\[ \phi(a) = \phi(a_0) + \phi(a_1) \epsilon_n + \cdots + \phi(a_n) \epsilon_n^n \]  

(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

\[ \phi(a \epsilon_n) = \phi(a) \epsilon_n \]  

(5.1.6)
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

$$(5.1.7)\quad \beta(a, b) = \sum_{j=0}^{n} \sum_{l=0}^{n} \beta(a_j, b_l) \epsilon_n^{j+l}$$

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

$$(5.1.8)\quad \beta(a, b \epsilon_n) = \beta(a, b) \epsilon_n$$

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

$$(5.1.9)\quad \beta(a, a) = 0$$

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

$$(5.1.11)\quad \beta(a, a) = \sum_{j=0}^{n} \sum_{l=0}^{n} \beta(a_j, a_l) \epsilon_n^{j+l},$$

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

$$(5.1.12)\quad \beta(a_j, a_l) + \beta(a_l, a_j) = 0$$

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
5.2. ADJOINING NILPOTENT ELEMENTS TO ALGEBRAS

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

\[
(5.2.1) \quad a \cdot b = \sum_{j=0}^{n} \sum_{l=0}^{n} a_j b_l \epsilon_n^{j+l}
\]
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, \(a \cdot b\) 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 \(C[\epsilon_n]\) 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 \((A, [\cdot, \cdot], \alpha)\) is a Lie algebra over \(k\), then \([\cdot, \cdot], \alpha\) 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

\[
(5.2.2) \quad [a, b]_{A[\epsilon_n]} = \sum_{j=0}^{n} \sum_{l=0}^{n} [a_j, b_l]_A \epsilon_n^{j+l}
\]
defines an element of \(A[\epsilon_n]\) as well. As in the previous section, \([a, a]_{A[\epsilon_n]} = 0\) for every \(a \in A[\epsilon_n]\), because of the analogous property of \([\cdot, \cdot], \alpha\) on \(A\). One can verify that (5.2.2) satisfies the Jacobi identity on \(A[\epsilon_n]\), using the Jacobi identity for \([\cdot, \cdot], \alpha\) 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 \(f(T) = \sum_{j=0}^{\infty} f_j T^j\) is an element of \(A[[T]]\), then

\[
(5.2.3) \quad f_0 + f_1 \epsilon_n + \cdots + f_n \epsilon_n^n
\]
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

$$(5.3.1) \quad a = a_0 + a_1 \epsilon_n + \cdots + a_n \epsilon_n^n,$$

where $a_0, a_1, \ldots, a_n \in A$. The mapping

$$(5.3.2) \quad a \mapsto a_0$$

defines an algebra homomorphism from $A[\epsilon_n]$ onto $A$. If $a$ is invertible in $A[\epsilon_n]$, then it follows that $a_0$ is invertible in $A$.

Note that

$$(5.3.3) \quad (e - a) \sum_{l=0}^{n} a^l = \left( \sum_{l=0}^{n} a^l \right) (e - a) = e - a^{n+1}$$

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

$$(5.3.4) \quad (e - a) \sum_{l=0}^{n} a^l = \left( \sum_{l=0}^{n} a^l \right) (e - a) = e.$$

This means that $e - a$ is invertible in $A[\epsilon_n]$, with

$$(5.3.5) \quad (e - a)^{-1} = \sum_{l=0}^{n} a^l.$$ 

If

$$(5.3.6) \quad b = b_0 + b_1 \epsilon_n + \cdots + b_n \epsilon_n^n$$

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 $a = e + \alpha$, $b = e + \beta$,

$$a b = e + \alpha + \beta + \alpha \beta,$$

and $\alpha$, $\beta$ are multiples of $\epsilon_n$. As in the preceding paragraph, $a$ and $b$ are invertible elements of $A[\epsilon_n]$, with

$$a^{-1} = \sum_{l=0}^{n} (-\alpha)^l = e + \sum_{l=1}^{n} (-\alpha)^l$$

and

$$b^{-1} = \sum_{l=0}^{n} (-\beta)^l = e + \sum_{l=1}^{n} (-\beta)^l.$$

We also have that

$$a b a^{-1} b^{-1} = e + a_1 b_1 \epsilon_n^2 - b_1 a_1 \epsilon_n^2 + O(\epsilon_n^3) = e + \alpha \beta - \beta \alpha + O(\epsilon_n^3),$$

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 $\alpha = a_1 \epsilon_1$, $\beta = b_1 \epsilon_1$, and (5.3.7) reduces to

$$a b = e + \alpha + \beta = e + a_1 \epsilon_1 + b_1 \epsilon_1.$$ 

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 $a b a^{-1} b^{-1} = e$.

If $n = 2$, then $\alpha = a_1 \epsilon_2 + a_2 \epsilon_2^2$, $\beta = b_1 \epsilon_2 + b_2 \epsilon_2^2$, and (5.3.7) reduces to

$$a b = e + a_1 \epsilon_2 + a_2 \epsilon_2^2 + b_1 \epsilon_2 + b_2 \epsilon_2^2 + a_1 b_1 \epsilon_2^2.$$ 

Similarly, (5.3.10) reduces to

$$a b a^{-1} b^{-1} = e + a_1 b_1 \epsilon_2^2 - b_1 a_1 \epsilon_2^2.$$ 

### 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_{0,1} \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 \quad \text{for every } a \in A[\epsilon_1, \eta_1],
\end{equation}
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 \quad \text{for every } a \in A[\epsilon_1, \eta_1] \text{ and } b \in B[\epsilon_1, \eta_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 \quad \text{for every } a \in A[\epsilon_1, \eta_1] \text{ and } b \in B[\epsilon_1, \eta_1].
\end{equation}

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,

$$a \mapsto a_{0,0}$$

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 $A$.

For every $a \in A[\epsilon_1, \eta_1]$, where $a'$ 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

$$a \mapsto a_{0,0}$$

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 $A$.

It follows that $e - a$ is invertible in $A[\epsilon_1, \eta_1]$, with

$$e - a \mapsto e - a$$

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 $A$.

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

$$a = a - e = a_{1,0} \epsilon_1 + a_{0,1} \eta_1 + a_{1,1} \epsilon_1 \eta_1$$

and

$$b = b - e = b_{1,0} \epsilon_1 + b_{0,1} \eta_1 + b_{1,1} \epsilon_1 \eta_1.$$
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
\[(5.5.11) \quad a = e + a_{1,0} \epsilon_1, \quad b = e + b_{0,1} \eta_1\]
and \(\alpha = a_{1,0} \epsilon_1, \beta = b_{0,1} \eta_1\). In this case, (5.5.8) reduces to
\[(5.5.12) \quad a b = e + a_{1,0} \epsilon_1 + b_{0,1} \eta_1 + a_{1,0} b_{0,1} \epsilon_1 \eta_1.\]
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
\[(5.5.15) \quad a b a^{-1} b^{-1} = e + a_{1,0} b_{0,1} \epsilon_1 \eta_1 - b_{0,1} a_{1,0} \epsilon_1 \eta_1.\]

### 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
\[(5.6.3) \quad h_n = \sum_{j=0}^{n} \beta(f_j, g_{n-j}).\]
5.6. **DIFFERENTIATION**

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

\[
(n + 1) \cdot h_{n+1} = \sum_{j=0}^{n} \beta((j + 1)f_{j+1}, g_{n-j}) + \sum_{j=0}^{n} \beta(f_j, (n-j+1)g_{n-j+1})
\]

for every \( n \geq 0 \). By the definition (5.6.3) of \( h_n \), we have that

\[
(n + 1) \cdot h_{n+1} = \sum_{j=0}^{n+1} (n + 1) \cdot \beta(f_j, g_{n+1-j})
\]

\[
= \sum_{j=0}^{n+1} j \cdot \beta(f_j, g_{n+1-j}) + \sum_{j=0}^{n+1} (n + 1 - j) \cdot \beta(f_j, g_{n+1-j})
\]

\[
= \sum_{j=1}^{n+1} j \cdot \beta(f_j, g_{n+1-j}) + \sum_{j=0}^{n} (n + 1 - j) \cdot \beta(f_j, g_{n+1-j})
\]

\[
= \sum_{j=1}^{n} (j + 1) \cdot \beta(f_{j+1}, g_{n-j}) + \sum_{j=0}^{n} (n-j+1) \cdot \beta(f_j, g_{n-j+1})
\]

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, \( f'(T) \in k[[T]] \), \( g'(T) \in A[[T]] \), 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

$$f(t) = \sum_{j=0}^{n} f_j t^j$$

(5.7.1)

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

$$f(T) \mapsto f(t)$$

(5.7.2)

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

$$h(T) = \beta(f(T), g(T))$$

(5.7.3)

can be defined as an element of $C[T]$ as in Section 4.5. Under these conditions, one can check that

$$h(t) = \beta(f(t), g(t))$$

(5.7.4)

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

$$h(t) = f(t) \cdot g(t)$$

(5.7.5)

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

$$\epsilon^2 = 0.$$  

(5.7.6)

Note that

$$(t + \epsilon)^j = t^j + j \cdot t^{j-1} \epsilon$$

(5.7.7)
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 \( a b \). 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 (5.8.15)
\[ (k[[X_1, \ldots, X_l]])[[Y_1, \ldots, Y_m]] \]
and (5.8.16)
\[ k[[X_1, \ldots, X_l, Y_1, \ldots, Y_m]] \]
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
\[ t^\alpha = t_1^{\alpha_1} \cdots t_n^{\alpha_n}. \]

Here $t_j^{\alpha_j}$ 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
\[ t^{\alpha + \beta} = t^\alpha t^\beta. \]

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
\[ f(t) = \sum_{|\alpha| \leq N} f_\alpha t^\alpha \]
defines an element of $A$, where $f_\alpha t^\alpha$ is defined using scalar multiplication on $A$ for each multi-index $\alpha$. The mapping
\[ f(T) \mapsto f(t) \]
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
\[ h(t) = f(t) g(t). \]

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 $f(T)$
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

\[
a_j a_l = a_l a_j
\]  

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

\[
a^\alpha = a_1^{\alpha_1} \cdots a_n^{\alpha_n},
\]

where \( a_j^{\alpha_j} \) is interpreted as being equal to \( e \) when \( \alpha_j = 0 \). Note that

\[
a^{\alpha + \beta} = a^{\alpha} a^{\beta}
\]

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

\[
f(a) = \sum_{|\alpha| \leq N} f_{\alpha} a^\alpha,
\]

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

\[
h(a) = f(a) g(a).
\]

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

\[
\alpha_j(l) = \alpha_j \quad \text{when } j \neq l
\]

and

\[
\alpha_l(l) = \begin{cases} 
\alpha_l - 1 & \text{when } \alpha_l \geq 1 \\
0 & \text{when } \alpha_l = 0.
\end{cases}
\]

Similarly, let \( \alpha^+(l) \) be the multi-index of length \( n \) defined by

\[
\alpha_j^+(l) = \begin{cases} 
\alpha_j & \text{when } j \neq l \\
\alpha_l + 1 & \text{when } j = l.
\end{cases}
\]
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

\[(5.10.4) \quad \frac{\partial f(T)}{\partial T_l} f(T) = \sum_{\alpha \in (\mathbb{Z}_+ \cup \{0\})^n} (\alpha_l + 1) \cdot f_{\alpha + (l)} T^\alpha.\]

This is basically the same as

\[(5.10.5) \quad \sum_{\alpha \in (\mathbb{Z}_+ \cup \{0\})^n} \alpha_l \cdot f_{\alpha} T^{\alpha(l)} = \sum_{\alpha_l \geq 1} \alpha_l \cdot f_{\alpha} T^{\alpha(l)},\]

where the second sum is taken over all multi-indices $\alpha$ with $\alpha_l \geq 1$. Note that

\[(5.10.6) \quad f(T) \mapsto \partial_l f(T)\]

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

\[(5.10.7) \quad \partial_l (\partial_m f(T)) = \partial_m (\partial_l f(T))\]

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

\[(5.10.8) \quad \partial_l (f(T) g(T)) = (\partial_l f(T)) g(T) + f(T) (\partial_l g(T)),\]

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 coordinatewise 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),
\[(5.10.10) \quad (t_l + u_l)^\alpha_l = t_l^\alpha_l + \alpha_l \cdot t_l^{\alpha_l - 1} u_l\]
for each $l = 1, \ldots, n$ when $\alpha_l \geq 1$, because $u^2_l = 0$. If $\alpha_l = 0$, then $(t_l + u_l)^\alpha_l = t_l^{\alpha_l} = 1$. Thus
\[(5.10.11) \quad (t_l + u_l)^\alpha_l = t_l^{\alpha_l} + \alpha_l \cdot t_l^{\alpha_l(l)} u_l\]
for every $l = 1, \ldots, n$. Using this, one can check that
\[(5.10.12) \quad (t + u)^\alpha = t^\alpha + \sum_{l=1}^{n} \alpha_l \cdot t^{\alpha(l)} u_l.\]
If $f(T) \in A[T_1, \ldots, T_n]$, then it follows that
\[(5.10.13) \quad f(t + u) = f(t) + \sum_{l=1}^{n} (\partial_l f)(t) u_l.\]

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),
\[(5.10.15) \quad (a_l + u_l)^\alpha_l = a_l^{\alpha_l} + \alpha_l \cdot a_l^{\alpha_l(l)} u_l\]
for every $l = 1, \ldots, n$. This implies that
\[(5.10.16) \quad (a + u)^\alpha = a^\alpha + \sum_{l=1}^{n} \alpha_l \cdot a^{\alpha(l)} u_l,\]
as before. If $f(T) \in k[T_1, \ldots, T_n]$, then we get that
\[(5.10.17) \quad f(a + u) = f(a) + \sum_{l=1}^{n} (\partial_l f)(a) u_l.\]
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 $\alpha$. The space of these formal differential operators can be defined as the space of functions $\alpha \mapsto a^\alpha(T)$ from $(\mathbb{Z}_+ \cup \{0\})^n$ into $k[[T_1, \ldots, T_n]]$ such that $a^\alpha(T) = 0$ for all but finitely many $\alpha$. 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 $k[[T_1, \ldots, T_n]]$ indexed by $(\mathbb{Z}_+ \cup \{0\})^n$. We can identify elements of $k[[T_1, \ldots, T_n]]$ with sums of the form (5.11.2) with $N = 0$.

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

$$\partial_l \partial^{\beta} = \partial^{\beta+\langle l \rangle},$$

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.
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

\[
\partial^\alpha f(T) = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f(T)
\]

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

\[
\left( \sum_{|\alpha| \leq N} a^\alpha(T) \partial^\alpha \right) f(T) = \sum_{|\alpha| \leq N} a^\alpha(T) \partial^\alpha f(T)
\]

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

\[
\text{Hom}_k(A[[T_1, \ldots, T_n]], A[[T_1, \ldots, T_n]])
\]

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.

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

\[
a(T) = (a^1(T), \ldots, a^n(T))
\]

be an $n$-tuple of formal power series in $T_1, \ldots, T_n$ with coefficients in $k$. Put

\[
D_n(T) = \sum_{j=1}^n a^j(T) \partial_j,
\]

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

\[
(5.12.7) \quad D_{a(T)} D_{b(T)} - D_{b(T)} D_{a(T)} = \sum_{j=1}^{n} \sum_{l=1}^{n} \left( a^j(T) \partial_j b^l(T) - b^l(T) \partial_j a^j(T) \right) \partial_l,
\]

using (5.11.3). Put \( c(T) = (c^1(T), \ldots, c^n(T)) \), where

\[
(5.12.4) \quad c^l(T) = \sum_{j=1}^{n} \left( a^j(T) \partial_j b^l(T) - b^l(T) \partial_j a^j(T) \right)
\]

for each \( l = 1, \ldots, n \). Thus \( c(T) \) is another \( n \)-tuple of elements of \( k[[T_1, \ldots, T_n]] \), and

\[
(5.12.5) \quad D_{a(T)} D_{b(T)} - D_{b(T)} D_{a(T)} = D_{c(T)}.
\]

If \( f(T) \in k[[T_1, \ldots, T_n]] \), then

\[
(5.12.6) \quad D_{a(T)} f(T) = \sum_{j=1}^{n} a^j(T) \partial_j f(T)
\]

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

\[
(5.12.7) \quad a^j(T) = \delta(T_j)
\]

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

\[
(5.12.8) \quad \delta(f(T)) = D_{a(T)} f(T),
\]

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 \( f(T) \in k[[T_1, \ldots, T_n]] \).
5.13. HOMOGENEOUS FORMAL POLYNOMIALS

Let us say that \( f(T) \in \sum_{\alpha \in (\mathbb{Z}_{c} \cup \{0\})^n} f_\alpha T^\alpha \in k[[T_1, \ldots, T_n]] \) 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

\[
(5.12.9) \quad g(T) T^\beta
\]

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 homogenous of degree \( d \) if \( f(T) \) can be expressed as

\[
(5.13.1) \quad f(T) = \sum_{|\alpha| = d} f_\alpha T^\alpha,
\]

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 \( \partial f(T) \) 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

\[
(5.13.2) \quad a^j(T) = \sum_{l=1}^{n} a^j_l T_l
\]

be elements of \( k[[T_1, \ldots, T_n]] \) for \( j = 1, \ldots, n \), so that \( a^j_l \in k \) for all \( j, l = 1, \ldots, n \). Thus \( \partial a^j(T) = a^j_l \) for every \( j, l = 1, \ldots, n \). Put \( a(T) = (a^1(T), \ldots, a^n(T)) \), and
120 \textbf{CHAPTER 5. SOME RELATED NOTIONS}

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

\begin{equation}
(5.13.3) \quad f(T) = \sum_{j=1}^{n} f_j T_j
\end{equation}

is homogeneous of degree 1, so that \( f_j \in k \) for \( j = 1, \ldots, n \), then

\begin{equation}
(5.13.4) \quad 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^j_l f_j T_l.
\end{equation}

Let

\begin{equation}
(5.13.5) \quad b^j(T) = \sum_{l=1}^{n} b^j_l T_l
\end{equation}

be elements of \( k_1[T_1, \ldots, T_n] \) for \( j = 1, \ldots, n \), and put \( b(T) = (b^1(T), \ldots, b^n(T)) \). Put

\begin{equation}
(5.13.6) \quad c^j(T) = \sum_{l=1}^{n} (a^j(T) b^j_l - b^j(T) a^j_l) = \sum_{l=1}^{n} \sum_{m=1}^{n} (a^j_m b^j_l - b^j_m a^j_l) T_m
\end{equation}

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 \( k_1[T_1, \ldots, T_n] \) 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

\begin{equation}
(5.13.7) \quad D_{a(T)} D_{b(T)} - D_{b(T)} D_{a(T)} = D_{c(T)},
\end{equation}

as in (5.12.5).

\section{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

\begin{equation}
(5.14.1) \quad L = \sum_{|\alpha| \leq N} a_{\alpha}(T) \partial^{\alpha}
\end{equation}

in \( \partial_1, \ldots, \partial_n \) with coefficients in \( k[T_1, \ldots, T_n] \). Let us say that \( L \) is \textit{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 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

$$\rho_a(v) = \beta(a, v), \quad (6.1.1)$$

where $a \in A$ and $v \in V$. The bilinearity 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

$$a \cdot v = \beta(a, v) \quad (6.1.2)$$

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 $\beta$ 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

$$N_V(\beta(a, v)) \leq C N_A(a) N_V(v)$$

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

$$N_V(\rho_a(v)) \leq C N_A(a) N_V(v)$$

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

$$\|\rho_a\|_{op,V,V} \leq C N_A(a).$$

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

$$\beta(a, w) \in W$$

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

$$\rho_a(W) \subseteq W$$

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 $a \in A$
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 \( a b \in A \) 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 \( \rho_a^V \) and \( \rho_a^Z \) 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

\[
(6.2.6) \quad \phi \circ \rho_a^V = \rho_a^Z \circ \phi
\]

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
6.3. QUOTIENT ACTIONS

\[ V/W \] can be defined as a module over \( k \) too, as in Section 2.11. Let \( q \) be the corresponding quotient mapping from \( V \) onto \( V/W \).

Suppose that \( \phi \) is a homomorphism from \( V \) into itself, as a module over \( k \), such that

\[ \phi(W) \subseteq W. \]  

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 \( V \) onto \( V/W \) such that

\[ \psi \circ q = q \circ \phi \]  

as mappings from \( V \) into \( W \), by standard arguments. Equivalently, this means that

\[ \psi(q(v)) = q(\phi(v)) \]  

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

\[ \beta^{V/W}(a, q(v)) = q(\beta^V(a, v)) \]  

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

\[ \rho_a^{V/W} \circ q = q \circ \rho_a^V, \]  

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 \( \text{Hom}_k(V/W, V/W) \). 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

\[ a \cdot q(v) = q(a \cdot v) \]  

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

\[ q(v) \cdot a = q(v \cdot a) \]  

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

\[
\rho_a \circ \rho_b = \rho_{a b}
\]

for every \( a, b \in A \), which is the same as saying that

\[
\rho_a(\rho_b(v)) = \rho_{a b}(v)
\]

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

\[
a \cdot (b \cdot v) = (a b) \cdot v
\]

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 (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

\[
(v \cdot a) \cdot b = v \cdot (a \cdot b)
\]

for every \( a, b \in A \) and \( v \in V \), then \( V \) is said to be a 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

\[
\rho_a(v) = v \cdot a
\]

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

\[
\rho_b(\rho_a(v)) = \rho_{ab}(v)
\]

for every \( a, b \in A \) and \( v \in V \), which means that

\[
\rho_b \circ \rho_a = \rho_{ab}
\]

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 \((A, \cdot, [\cdot, \cdot]_A)\) 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 \( a b \). 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

\[
(6.5.5) \quad \rho_{[a,b]} = \rho_{ab} - \rho_{ba} = \rho_b \circ \rho_a - \rho_a \circ \rho_b.
\]

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

\begin{equation}
\rho_a(W) \subseteq W
\end{equation}

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

\begin{equation}
a \cdot w \in W
\end{equation}

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

\begin{equation}
w \cdot a \in W
\end{equation}

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

\begin{equation}
a(W) \subseteq W
\end{equation}

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

$$\phi \circ \rho^V_a = \rho^W_a \circ \phi$$

for every $a \in A$, which is the same as saying that

$$\phi(\rho^V_a(v)) = \rho^W_a(\phi(v))$$

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

$$a \cdot v = \rho^V_a(v), \quad a \cdot w = \rho^W_a(w)$$

for every $a \in A, v \in V$, and $w \in W$. Thus (6.7.2) may be reexpressed as

$$\phi(a \cdot v) = a \cdot \phi(v)$$

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 $A$.

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 $\rho^V_{a/W}$ from $V/W$ into itself such that

$$\rho^V_{a/W} \circ q = q \circ \rho^V_a$$

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 $A$. 

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
\[
q(v) \cdot a = q(v \cdot a)
\]
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 \( j \in I \), \( \beta^V_j \) 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
\[
(\beta(a, v))_j = \beta^V_j(a, v_j)
\]
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
\[
\beta(a, v) \in \bigoplus_{j \in I} V_j
\]
for every \( a \in A \).

Equivalently, if the bilinear action of \( A \) on \( V_j \) is given by \( \rho^V_j \) 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
\[
(\rho(a)(v))_j = \rho^V_j(a, v_j).
\]
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
\[
(a \cdot v)_j = a \cdot v_j
\]
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

\[
(v \cdot a)_j = v_j \cdot a
\]

for each \( j \in I \).

Let \( A \) be an associative algebra over \( k \). If \( \rho^{V_j} \) 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 \( \prod_{j \in I} V_j \) 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 \( \rho^{V_j} \) 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

\[
V_1 + V_2 = \{v_1 + v_2 : v_1 \in V_1, \ v_2 \in V_2\}
\]

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

\[
(v_1, v_2) \mapsto v_1 + v_2
\]

defines a homomorphism from \( V_1 \oplus V_2 \) onto \( V_1 + V_2 \), as modules over \( k \). If

\[
V_1 \cap V_2 = \{0\},
\]

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
\[ \rho_a \text{ is a module homomorphism from } V \text{ into itself for every } a \in A, \text{ and } a \mapsto \rho_a \text{ is a module homomorphism from } A \text{ into } \text{Hom}_k(V, V), \text{ as in Section 6.1. Let } W \text{ be another module over } k, \text{ and let } \mu \text{ be a mapping from } V \times V \text{ into } W \text{ that is bilinear over } k. \text{ Suppose that for every } a \in A \text{ there is an } \bar{a} \in A \text{ such that} \]

\[ (6.10.1) \quad \mu(\rho_a(u), v) = \mu(u, \rho_{\bar{a}}(v)) \]

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

\[ (6.10.2) \quad (V_0)^\perp = (V_0)^{\perp \mu} = \{ u \in V : \text{ for every } v \in V_0, \mu(u, v) = 0 \}. \]

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 \( v \in V_0 \) be given, and let \( \bar{a} \in A \) be as in (6.10.1). Observe that

\[ (6.10.4) \quad \mu(\rho_a(u), v) = \mu(u, \rho_{\bar{a}}(v)) = 0, \]

because \( \rho_{\bar{a}}(v) \in V_0 \), 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 \( \bar{a} = a^* \), 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

\[ (6.10.9) \quad (\rho_a)^* = \rho_{a^*} \]
6.11. REPRESENTATIONS AND FORMAL POWER SERIES

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 \((A, [\cdot, \cdot]_A)\) 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 \( \rho \) with \( \theta \) is that \( \rho_a \) be antisymmetric with respect to \( \theta \) for every \( a \in A \), so that

\[
\theta(\rho_a(u), v) = -\theta(u, \rho_a(v))
\]

(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
$k[[T]]$, as in Section 4.6. 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
\begin{align}
(6.11.3) \quad \text{Hom}_{k[T]}(V[T], V[T])
\end{align}

and a natural injective homomorphism from
\begin{align}
(6.11.4) \quad (\text{Hom}_k(V, V))[T]
\end{align}

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 4.6. 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 $a \cdot b$. The corresponding \textit{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 $b \cdot a$ 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 $a, b \in M_n(A)$ be matrices 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

$$c_{j,r} = \sum_{l=1}^{n} a_{j,l} b_{l,r}$$  \hspace{1cm} (6.12.1)

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

$$\bar{c}_{j,r} = \sum_{l=1}^{n} b_{l,r} a_{j,l}$$  \hspace{1cm} (6.12.2)

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\textit{ transposes} of $a, b,$ and $c$, respectively, as in Section 2.8. Thus

$$c^t_{j,r} = c_{r,j} = \sum_{l=1}^{n} a_{r,l} b_{l,j}$$  \hspace{1cm} (6.12.3)

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

$$a \mapsto a^t$$  \hspace{1cm} (6.12.4)

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 $M_n(A^{op})$. 

6.12. OPPOSITE ALGEBRAS 137
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

\begin{equation}
(6.13.1)
a \cdot x = (ax_1, \ldots, ax_n)
\end{equation}

and

\begin{equation}
(6.13.2)
x \cdot a = (x_1a, \ldots, x_na),
\end{equation}

respectively. If $A$ is commutative, then (6.13.1) and (6.13.2) are the same. Note that

\begin{equation}
(6.13.3)
(a \cdot x) \cdot b = a \cdot (x \cdot b)
\end{equation}

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

\begin{equation}
(6.13.4)
(T^L_\alpha(x))_j = \sum_{l=1}^{n} \alpha_{j,l} x_l
\end{equation}

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

\begin{equation}
(6.13.5)
(T^R_\alpha(x))_j = \sum_{l=1}^{n} x_l \alpha_{j,l}
\end{equation}

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

\begin{equation}
(6.13.6)
(T^L_\alpha(x)) \cdot a = T^L_\alpha(a \cdot x)
\end{equation}

and

\begin{equation}
(6.13.7)
a \cdot (T^R_\alpha(x)) = T^R_\alpha(a \cdot x)
\end{equation}

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

\begin{equation}
(6.13.8)
T^L_{t \alpha}(x) = t T^L_\alpha(x)
\end{equation}

and

\begin{equation}
(6.13.9)
T^R_{t \alpha}(x) = t T^R_\alpha(x)
\end{equation}
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,

$$T^L_{\alpha+\beta}(x) = T^L_\alpha(x) + T^L_\beta(x)$$

and

$$T^R_{\alpha+\beta}(x) = T^R_\alpha(x) + T^R_\beta(x)$$

for every $x \in A^n$. One can check that

$$T^L_\alpha(T^L_\beta(x)) = T^L_{\alpha \beta}(x)$$

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

$$T^R_\alpha(T^R_\beta(x)) = T^R_\gamma(x)$$

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 $T^L_\alpha$ and $T^R_\alpha$ are equal to the identity mapping on $A^n$. Let $u^1, \ldots, u^n$ be the elements of $A^n$ with $u^l_j = e$ when $j = l$ and $u^l_j = 0$ when $j \neq l$. Thus

$$x = \sum_{l=1}^n x_l \cdot u^l = \sum_{l=1}^n u^l \cdot x_l$$

for every $x \in A^n$.

If $T$ is any homomorphism from $A^n$ into itself, as a right module over $A$, then

$$T(x) = T\left(\sum_{l=1}^n u^l \cdot x_l\right) = \sum_{l=1}^n T(u^l) \cdot x_l$$

for every $x \in A^n$. This means that $T$ can be represented in a unique way as $T^L_\alpha$, 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

$$T(x) = T\left(\sum_{l=1}^n x_l \cdot u^l\right) = \sum_{l=1}^n x_l \cdot T(u^l)$$

for every $x \in A^n$. This implies that $T$ can be represented in a unique way as $T^R_\alpha$, 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 $W \neq \{0\}$, $V$

\begin{equation}
(6.14.1) \quad \rho_a(W) \subseteq W
\end{equation}

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

\begin{equation}
(6.14.2) \quad \phi(\rho^1_a(v_1)) = \rho^2_a(\phi(v_1)) = 0,
\end{equation}

so that $\rho^1_a(v_1)$ 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,

\begin{equation}
(6.14.3) \quad \rho^2_a(\phi(v_1)) = \phi(\rho^1_a(V_1)) \subseteq \phi(V_1)
\end{equation}

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 spaces 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 $\mathcal{L}^\rho(V)$ of $\phi \in \mathcal{L}(V)$ that intertwine $\rho$. It is easy to see that $\mathcal{L}^\rho(V)$ is a subalgebra of $\mathcal{L}(V)$, and that $\mathcal{L}^\rho(V)$ contains the identity mapping on $V$. If $\rho$ is irreducible, then every nonzero element of $\mathcal{L}^\rho(V)$ 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 $a_1 b_1 \in A_1$ and $a_2 b_2 \in A_2$ are expressed as $a_1 b_1$ and $a_2 b_2$, respectively, then

\[(a_1, 0)(0, a_2) = 0\]  

(7.1.1)

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

\[A_1 + A_2 = \{a_1 + a_2 : a_1 \in A_1, a_2 \in A_2\}\]

(7.1.2)

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

\[a_1 a_2 = a_2 a_1 = 0\]

(7.1.3)

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,

\[(a_1, a_2) \mapsto a_1 + a_2\]

(7.1.4)
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

(7.1.5) $[a_1, a_2]_A = 0$

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

(7.1.6) $\phi(x) \psi(y) = \psi(y) \phi(x) = 0$

for every $x, y \in A$. This implies that

(7.1.7) $(\phi(x) + \psi(x)) (\phi(y) + \psi(y)) = \phi(x) \phi(y) + \phi(x) \psi(y) + \psi(x) \phi(y) + \psi(x) \psi(y)$

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

(7.1.8) $[\phi(x), \psi(y)]_B = 0$

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

(7.1.9) $\rho^1_a \circ \rho^2_b = \rho^2_b \circ \rho^1_a$

for every $a, b \in A$. Under these conditions,

(7.1.10) $(\rho^1 + \rho^2)_a = \rho^1_a + \rho^2_a$

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

(7.2.1) $\tilde{T}(\phi) = \phi \circ T$

defines another module homomorphism from $V$ into $W$. This defines $\tilde{T}$ as a homomorphism from $\text{Hom}_k(V, W)$ into itself, as a module over $k$. 
7.3. MULTILINEAR MAPPINGS

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, $\tilde{R}$ 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

$$(7.2.2) \quad (T \circ R)(\phi) = \phi \circ (T \circ R) = (\phi \circ T) \circ R = \tilde{R}(\phi).$$

This means that

$$(7.2.3) \quad (T \circ R) = \tilde{R} \circ T$$

as mappings from $\text{Hom}_k(V, W)$. Note that

$$(7.2.4) \quad T \mapsto \tilde{T}$$

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 $\tilde{T}$ 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 $V$, so that $\tilde{\rho}_a$ 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 $\tilde{\rho}_a$ is an opposite algebra homomorphism from $A$ into the algebra of module homomorphisms from $\text{Hom}_k(V, W)$ into itself. Similarly, if $a \mapsto \rho_a$ is an opposite algebra homomorphism from $A$ into $\text{Hom}_k(V, V)$, then $a \mapsto \tilde{\rho}_a$ 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, $\tilde{\rho}_a$ is defined as a module homomorphism from $V$ into $W$ for every $a \in A$. Under these conditions, one can check that $a \mapsto -\tilde{\rho}_a$ 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

\[
(7.3.1) \quad \mu(v_1, \ldots, v_{j-1}, v_l, v_{j+1}, \ldots, v_l, v_1, v_2, \ldots, v_n) = -\mu(v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_l, v_1, v_2, \ldots, v_n)
\]

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

\[
(7.3.2) \quad \mu(v_1, \ldots, v_n) = 0
\]

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

\[
(7.3.3) \quad (A_j(\mu))(v_1, \ldots, v_n) = \mu(v_1, \ldots, v_{j-1}, A_j(v_j), v_{j+1}, \ldots, v_n).
\]

This defines a mapping $A_j(\mu)$ from $\prod_{j=1}^n V_j$ into $W$, which corresponds to composing $\mu$ with $A_j$ in the $j$th variable. Note that $A_j(\mu)$ is multilinear over $k$, because $A_j$ is linear over $k$. It is easy to see that $A_j$ is linear over $k$, as a mapping from $L(V_1, \ldots, V_n; W)$ into itself. The mapping

\[
(7.3.4) \quad A_j \mapsto A_j
\]

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 $B_j \circ A_j$ of $A_j$ 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

\[
(7.3.5) \quad ((B_j \circ A_j)(\mu))(v_1, \ldots, v_n)
\]

\[
= \mu(v_1, \ldots, v_{j-1}, (B_j \circ A_j)(v_j), v_{j+1}, \ldots, v_n)
\]

\[
= \mu(v_1, \ldots, v_{j-1}, B_j(A_j(v_j)), v_{j+1}, \ldots, v_n).
\]
This is the same as

$$(7.3.6) \quad (\tilde{B}_j(\mu))(v_1, \ldots, v_{j-1}, A_j(v_j), v_{j+1}, \ldots, v_n) = (\tilde{A}_j(\tilde{B}_j(\mu)))(v_1, \ldots, v_n),$$

so that

$$(7.3.7) \quad (B_j \circ A_j) = \tilde{A}_j \circ \tilde{B}_j$$

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,

$$(7.4.4) \quad T_\mu(v_n) = \mu_{v_n}$$
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

\[ N_W(\mu_{v_n}(v_1, \ldots, v_{n-1})) = N_W(\mu(v_1, \ldots, v_n)) \leq \|\mu\|_{v_1, \ldots, v_n; W} N_{V_1}(v_1) \cdots N_{V_{n-1}}(v_{n-1}) N_{V_n}(v_n) \]

(7.4.5)

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

\[ \|\mu_{v_n}\|_{v_1, \ldots, v_{n-1}; W} \leq \|\mu\|_{v_1, \ldots, v_n; W} N_{V_n}(v_n). \]

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

\[ \|\mu_{v_n}\|_{v_1, \ldots, v_{n-1}; W} \leq \|T_\mu\|_{op} N_{V_n}(v_n). \]

(7.4.6)

If $v_1 \in V_1, \ldots, v_{n-1} \in V_{n-1}$, then we get that

\[ N_W(\mu(v_1, \ldots, v_{n-1}, v_n)) = N_W(\mu_{v_n}(v_1, \ldots, v_{n-1})) \leq \|\mu_{v_n}\|_{v_1, \ldots, v_{n-1}; W} N_{V_1}(v_1) \cdots N_{V_{n-1}}(v_{n-1}) \leq \|T_\mu\|_{op} N_{V_1}(v_1) \cdots N_{V_{n-1}}(v_{n-1}) N_{V_n}(v_n). \]

This implies that $\mu \in BL(V_1, \ldots, V_n; W)$, with

\[ \|\mu\|_{v_1, \ldots, v_n; W} \leq \|T_\mu\|_{op}. \]

(7.4.7)

It follows that

\[ \|T_\mu\|_{op} = \|\mu\|_{v_1, \ldots, v_n; W}, \]

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 $A_j(\mu)$ 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

\[ N_W((A_j(\mu))(v_1, \ldots, v_n)) = \|\mu\|_{v_1, \ldots, v_{j-1}, A_j(v_j), v_{j+1}, v_n} \leq \|\mu\|_{v_1, \ldots, v_n; W} N_{V_1}(v_1) \cdots N_{V_{j-1}}(v_{j-1}) N_{V_j}(A_j(v_j)) N_{V_{j+1}}(v_{j+1}) \cdots N_{V_n}(v_n) \]

(7.4.12)

\[ \leq \|A_j\|_{op, V_j} \|\mu\|_{v_1, \ldots, v_n; W} N_{V_1}(v_1) \cdots N_{V_n}(v_n). \]
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 \( \mu \in L(V_1, \ldots, V_n; W) \), then

\[
(7.5.2) \quad \mu(v_1, \ldots, v_{j-1}, a \cdot v_j, v_{j+1}, \ldots, v_n)
\]

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
\[
(a \cdot \mu)(v_1, \ldots, v_n) = a \cdot (\mu(v_1, \ldots, v_n)) + \sum_{j=1}^n (-\mu(v_1, \ldots, v_{j-1}, a \cdot v_j, v_{j+1}, \ldots, v_n)).
\]
(7.5.4)

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 \( ab \). 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
\[
ax = xa
\]
(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 \( (A, [\cdot, \cdot]_A) \) be a Lie algebra over \( k \). The centralizer of a set \( E \subseteq A \) in \( A \) as a Lie algebra is given by
\[
C_A(E) = \{ a \in A : [a, x]_A = 0 \text{ for every } x \in E \}.
\]
(7.6.2)
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

\[(7.6.3)\quad Z(A) = C_A(A) = \{a \in A : [a, x]_A = 0 \text{ for every } x \in A\},\]

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 $[x, y] = x y - y x$. Of course, (7.6.1) is the same as saying that $[a, x] = 0$. In this case, the centralizer of $E \subseteq A$ 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 $(A, [\cdot, \cdot]_A)$ be a Lie algebra over $k$ again, and 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

\[(7.6.4)\quad \rho_a(v) = 0\]

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

\[(7.6.5)\quad a \cdot v = 0\]

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 $\rho_a^V$ and $\rho_a^W$ are homomorphisms from $V$ and $W$ into themselves, as modules over $k$. If $\phi \in \text{Hom}_k(V, W)$, then it follows that $\phi \circ \rho_a^V$ and $\rho_a^W \circ \phi$ are homomorphisms from $V$ into $W$ as well, as modules over $k$. The mappings

$$\phi \mapsto \rho_a^W \circ \phi$$

and

$$\phi \mapsto -\phi \circ \rho_a^V$$

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

$$\rho_a(\phi) = \rho_a^W \circ \phi - \phi \circ \rho_a^V$$

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 representation 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 $\rho_a^V$ is a homomorphism from $V$ into itself, as a module over $k$, and

$$\beta(\rho_a^V(u), v), \beta(u, \rho_a^V(v))$$

define elements of $L_k(V, V; W)$. The mappings

$$\beta(u, v) \mapsto -\beta(\rho_a^V(u), v)$$

and

$$\beta(u, v) \mapsto -\beta(u, \rho_a^V(v))$$

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

$$(\rho_a(\beta))(u, v) = -\beta(\rho_a^V(u), v) - \beta(u, \rho_a^V(v))$$

defines a representation of $A$ as a Lie algebra over $k$ on $L_k(V, V; W)$ too. The condition that

$$\rho_a(\beta) = 0$$

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 $\rho_a^V$ 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 $\rho_a^V$ 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 $V_l^a$ be the set of $v_l \in V_l$ such that

\begin{equation}
\mu(v_1, \ldots, v_{l-1}, v_l, v_{l+1}, \ldots, v_n) = 0
\end{equation}

for every $v_j \in V_j$ with $1 \leq j \leq n$ and $j \neq l$. Note that $V_l^a$ 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 $v_l \in V_l^a$, then it is easy to see that $a \cdot v_l \in V_l^a$ too. This means that $V_l^a$ 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

\begin{equation}
(a \cdot \beta)(v, w) = a \cdot (\beta(v, w)) - \beta(a \cdot v, w) - \beta(v, a \cdot w)
\end{equation}

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

\begin{equation}
\text{tr}(ab) = \text{tr}(ba)
\end{equation}

for every $a, b \in M_n(A_0)$. Put

\begin{equation}
B_0(a, b) = \text{tr}(ab)
\end{equation}

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

\begin{equation}
B_0(ax, b) = \text{tr}((ax)b) = \text{tr}(a(xb)) = B_0(a, xb).
\end{equation}

We also have that

\begin{equation}
B_0(xa, b) = \text{tr}((xa)b) = \text{tr}(x(ab)) = \text{tr}(a(bx)) = B_0(a, bx).
\end{equation}
Of course, 
(7.8.5) \[ a \mapsto a x, \quad a \mapsto x a \]
define homomorphisms from \( M_n(A_0) \) into itself, as a module over \( k \). Similarly, 
(7.8.6) \[ C_x(a) = [x,a] = x a - a x \]
defines \( C_x \) as a homomorphism from \( M_n(A_0) \) into itself, as a module over \( k \). It is easy to see that 
(7.8.7) \[ B_0(C_x(a), b) = -B_0(a, C_x(b)) \]
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 
(7.8.8) \[ (T_a(v))_j = \sum_{l=1}^{n} a_{j,l} v_l \]
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 
(7.8.9) \[ \text{tr}(R \circ T) = \text{tr}(T \circ R), \]
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 
(7.8.10) \[ B(T_1, T_2) = \text{tr}(T_1 \circ T_2) \]
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 
(7.8.11) \[ C_R(T) = [R,T] = R \circ T - T \circ R \]
defines a homomorphism from \( \text{Hom}_k(V,V) \) into itself, as a module over \( k \). As in (7.8.7), we have that 
(7.8.12) \[ B(C_R(T_1), T_2) = -B(T_1, C_R(T_2)) \]
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 $V$. 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

\[ B_\rho([w, x]_A, y) = \text{tr}_V (\rho_{[w, x]} \circ \rho_y) = \text{tr}_V ([\rho_w, \rho_x] \circ \rho_y). \]  

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

\[ B_\rho(\text{ad}_w(x), y) = -B_\rho(x, \text{ad}_w(y)) \]  

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 \( \mathcal{A}_W \) be the collection of all \( T \in \text{Hom}_k(V, V) \) such that

\[
T(W) \subseteq W.
\]

(7.10.1)

Note that \( \mathcal{A}_W \) is a subalgebra of \( \text{Hom}_k(V, V) \). If \( T \in \mathcal{A}_W \), then let \( T_W \) be the restriction of \( T \) to \( W \), which defines a module homomorphism from \( W \) into itself. Of course,

\[
T \mapsto T_W
\]

(7.10.2)

defines an algebra homomorphism from \( \mathcal{A}_W \) 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 \( T \in \mathcal{A}_W \), 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

\[
T \mapsto T_{V/W}
\]

(7.10.4)

defines an algebra homomorphism from \( \mathcal{A}_W \) 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 \( V \) 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 \( T \in \mathcal{A}_W \), 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

\[
\text{tr}_V T = \text{tr}_W T_W + \text{tr}_{V/W} T_{V/W},
\]

(7.10.5)

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

\[ \text{tr}_V T = \text{tr}_W T_W. \]  

(7.10.7)

If $T_1, T_2 \in \text{Hom}_k(V, V)$, then put

\[ B_V(T_1, T_2) = \text{tr}_V (T_1 \circ T_2), \]  

(7.10.8)

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 $T_1, T_2 \in A_W$, and let $T_{1,W}, T_{2,W} \in \text{Hom}_k(W, W)$ and $T_{1,V/W}, T_{2,V/W} \in \text{Hom}_k(V/W, V/W)$ be as before. Note that

\[ (T_1 \circ T_2)_{W} = T_{1,W} \circ T_{2,W}, \quad (T_1 \circ T_2)_{V/W} = T_{1,V/W} \circ T_{2,V/W}, \]  

(7.10.9)

because (7.10.2) and (7.10.4) are algebra homomorphisms. It follows that

\[ B_V(T_1, T_2) = B_W(T_{1,W}, T_{2,W}) + B_{V/W}(T_{1,V/W}, T_{2,V/W}), \]  

(7.10.10)

by applying (7.10.5) to $T = T_1 \circ T_2$.

Let $(A, [\cdot, \cdot])$ be a Lie algebra over $k$, and let $B$ be an ideal in $A$, so that the quotient $A/B$ is defined as a Lie algebra over $k$ too. If $x \in A$, then $\text{ad}_x = \text{ad}_{A,x}$ is defined as a module homomorphism from $A$ into itself by

\[ \text{ad}_x(z) = \text{ad}_{A,x}(z) = [x, z] \]  

(7.10.11)

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

\[ \text{tr}_A(\text{ad}_x \circ \text{ad}_y) = \text{tr}_B((\text{ad}_x)_B \circ (\text{ad}_y)_B) + \text{tr}_{A/B}((\text{ad}_x)_{A/B} \circ (\text{ad}_y)_{A/B}), \]  

(7.10.12)

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

\[ V^\beta = \{ u \in V : \beta(u, v) = 0 \text{ for every } v \in V \} \]  

(7.11.1)

is a submodule of $V$. This may be called the radical of $\beta$ in $V$, as on p22 of [13].

Let $(A, [\cdot, \cdot])$ 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

\[
\beta(\text{ad}_w(x), y) = -\beta(x, \text{ad}_w(y))
\]

(7.11.2)

for every \( w, x, y \in A \), which is the same as saying that

\[
\beta([x, w], y) = \beta(x, [w, y])
\]

(7.11.3)

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

\[
\beta(x, y) = \text{tr}_A(\text{ad}_x \circ \text{ad}_y)
\]

(7.11.4)

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

\[
\text{tr}_A(\text{ad}_x \circ \text{ad}_y) = \text{tr}_B((\text{ad}_x)_B \circ (\text{ad}_y)_B)
\]

(7.11.5)

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

\[
\beta(x, y) = \text{tr}_A(\text{ad}_x \circ \text{ad}_y) = 0
\]

(7.11.6)

for every \( x \in B \) and \( y \in A \). This means that

\[
B \subseteq A^{\beta}
\]

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 $\bigotimes_{j=1}^n V_j$ 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 $v_1 \otimes \cdots \otimes v_n$. This means that every element of $\bigotimes_{j=1}^n V_j$ is the same as the module homomorphism $\phi$ from $\prod_{j=1}^n V_j$ to $W$ 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

$$(7.12.1) \quad \tilde{\mu}(v_1 \otimes \cdots \otimes v_n) = \mu(v_1, \ldots, v_n)$$

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

$$(7.12.2) \quad \phi_1(v_1) \otimes \cdots \otimes \phi_n(v_n).$$

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

$$(7.12.3) \quad \phi(v_1 \otimes \cdots \otimes v_n)$$

is equal to (7.12.2) 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.

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

\begin{equation}
(7.13.1) \quad \phi_x(w) = \phi(\delta_x w)
\end{equation}

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 $\text{Hom}_k(c_{00}(X, W), Z)$ and...
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 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 $\bigotimes_{\sigma(j)=1}^n V_{\sigma(j)}$ 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 $\bigotimes_{\sigma(j)=1}^n V_{\sigma(j)}$ with
\[
(7.14.1) \quad v_1 \otimes \cdots \otimes v_n \mapsto v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}
\]
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 $\bigotimes_{\sigma(j)=1}^n V_{\sigma(j)}$ defined by
\[
(7.14.2) \quad (v_1, \ldots, v_n) \mapsto v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)},
\]
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 $\bigotimes_{\tau(\sigma(j))=1}^n V_{\tau(\sigma(j))}$ with
\[
(7.14.3) \quad v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)} \mapsto v_{\tau(\sigma(1))} \otimes \cdots \otimes v_{\tau(\sigma(n))}
\]
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 (7.14.1) is a module homomorphism from \(\bigotimes_{j=1}^n V_j\) into \(\bigotimes V_{\tau(\sigma(j))}\) with
\[
(7.14.4) \quad v_1 \otimes \cdots \otimes v_n \mapsto v_{\tau(1)}(\sigma(1)) \otimes \cdots \otimes v_{\tau(n)}(\sigma(n))
\]
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 \(\bigotimes V_{\tau(\sigma(j))}\) with
\[
(7.14.5) \quad (v_1 \otimes \cdots \otimes v_n) \otimes (v_{n_1+1} \otimes \cdots \otimes v_n)
\]
is defined as a module over \(k\) as well. It is well known that there is a natural isomorphism between \(\bigotimes_{j=1}^n V_j\) and (7.14.5), 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
\[
(7.14.6) \quad (v_1 \otimes \cdots \otimes v_n) \otimes (v_{n_1+1} \otimes \cdots \otimes v_n)
\]
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,
\[
(7.14.7) \quad T^n V = \bigotimes_{j=1}^n V_j
\]
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 \(k\).
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 $j \mapsto f_j$ 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

\[ f(T) = \sum_{j > -\infty} f_j T^j, \]

to indicate that $f_j = 0$ 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 $(A[[T]]) T^n$ 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 $(A[[T]]) T^n$
correspond to formal power series in $T$ with coefficients in $A$ that vanish to
order $n - 1$, as in Section 4.3. Clearly $(A[[T]]) T^n$ is a submodule of $A((T))$ for
every $n \in \mathbb{Z}$, with

\begin{equation}
(A[[T]]) T^n \subseteq (A[[T]]) T^{n+1}
\end{equation}

and

\begin{equation}
A((T)) = \bigcup_{n=\infty}^{\infty} (A[[T]]) T^n.
\end{equation}

Let us say that the elements of a subset $E$ of $A((T))$ have poles of bounded
order if $E \subseteq (A[[T]]) T^n$ 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 sec-
tion. Let $l_0 \in \mathbb{Z}$ be given, and let

\begin{equation}
f_l(T) = \sum_{j >> -\infty} f_{l,j} T^j
\end{equation}

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

\begin{equation}
f_{l,j} = f_j
\end{equation}

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 $\{g_l(T)\}_{l=0}^{\infty}$ of elements of $A((T))$. If $\{f_l(T)\}_{l=0}^{\infty}$ and $\{g_l(T)\}_{l=0}^{\infty}$ are termwise eventually constant, then so are $\{\alpha f_l(T)\}_{l=0}^{\infty}$, $\{f_l(T) T^{r}\}_{l=0}^{\infty}$, and $\{f_l(T) + g_l(T)\}_{l=0}^{\infty}$. If the $f_l(T)$’s and $g_l(T)$’s have poles of bounded order, then the $\alpha f_l(T)$’s and $f_l(T) + g_l(T)$’s have the same property. In this case, if $r_0 \in \mathbb{Z}$, then the $f_l(T) T^{r}$’s have poles of bounded order for $l \geq l_0$ and $r \geq r_0$. If $\{f_l(T)\}_{l=0}^{\infty}$ and $\{g_l(T)\}_{l=0}^{\infty}$ eventually agree with $f(T), g(T) \in A((T))$ termwise, respectively, then $\{\alpha f_l(T)\}_{l=0}^{\infty}$, $\{f_l(T) T^{r}\}_{l=0}^{\infty}$, and $\{f_l(T) + g_l(T)\}_{l=0}^{\infty}$ eventually agree with $\alpha f(T)$, $f(T) T^{r}$, and $f(T) + g(T)$ termwise, respectively.

Let
\begin{equation}
a_l(T) = \sum_{j \geq -\infty} a_{l,j} T^j
\end{equation}
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
\begin{equation}
\sum_{l=0}^{n} a_l(T)
\end{equation}
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
\begin{equation}
\sum_{l=0}^{\infty} a_l(T).
\end{equation}

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
\begin{equation}
\sum_{l=0}^{\infty} \alpha a_l(T) = \alpha \sum_{l=0}^{\infty} a_l(T).
\end{equation}

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
\begin{equation}
\sum_{l=0}^{\infty} a_l(T) T^{r} = \left( \sum_{l=0}^{\infty} a_l(T) \right) T^{r}.
\end{equation}

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
\begin{equation}
\sum_{l=0}^{\infty} (a_l(T) + b_l(T)) = \sum_{l=0}^{\infty} a_l(T) + \sum_{l=0}^{\infty} b_l(T)
\end{equation}
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

\[
\phi(a) = \sum_{j=-\infty}^{\infty} \phi_j(a) T^j,
\]

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

\[
\phi(T) = \sum_{l=l_0}^{\infty} \phi_l T^l
\]

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

\[
(\phi(T))(a) = \sum_{l=l_0}^{\infty} \phi_l(a) T^l
\]

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

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

\[
a(T) = \sum_{m=m_0}^{\infty} a_m T^m
\]

be an element of \( A((T)) \), where \( m_0 \in \mathbb{Z} \). Thus

\[
(\phi(T))(a_m) = \sum_{l=l_0}^{\infty} \phi_l(a_m) T^l
\]

defines an element of \( B((T)) \) for every \( m \in \mathbb{Z} \), as in (8.3.3). Put

\[
(\phi(T))(a(T)) = \sum_{m=m_0}^{\infty} (\phi(T))(a_m) T^m,
\]

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

\[
(\phi(T))(a(T) T^r) = (\phi(T))(a(T)) T^r
\]

for every \( a(T) \in A((T)) \) and \( r \in \mathbb{Z} \). We also have that

\[
(\phi(T))(A[[T]]) T^n \subseteq (B[[T]]) T^{n+l_0}
\]

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

\[
(\phi(T) T^r)(a(T)) = (\phi(T))(a(T)) T^r
\]

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

\[
\phi_l(a_{n-l}) = 0
\]
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

\[
(8.4.9) \quad (\phi(T))(a(T)) = \sum_{n=0}^{\infty} ((\phi(T))(a(T)))_n T^n
\]
as an element of \( B(\langle T \rangle) \). One can check that this is equivalent to (8.4.3).

Let \( C \) be a third module over \( k \), and let \( \psi \) be a homomorphism from \( \langle T \rangle \) 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 \( (\text{Hom}_k(A,C))(\langle T \rangle) \). One can verify that the module homomorphism from \( A(\langle T \rangle) \) into \( C(\langle T \rangle) \) corresponding to (8.4.13) as before is the same as the composition of the homomorphisms from \( A(\langle T \rangle) \) into \( B(\langle T \rangle) \) and from \( B(\langle T \rangle) \) into \( C(\langle T \rangle) \) corresponding to \( \phi(T) \) and \( \psi(T) \), respectively.

### 8.5 Homomorphisms from \( A(\langle T \rangle) \) into \( B(\langle T \rangle) \)

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 \( A(\langle T \rangle) \) into \( B(\langle T \rangle) \), 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
\[
\phi(f(T) T^r) = \phi(f(T)) T^r
\]
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 $\{a_l(T)\}_{l=l_0}^\infty$ 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 $\{a_l(T)\}_{l=l_0}^\infty$ 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 $a \in A$ and $b \in B$, then $\beta(a, b)$ can be expressed as

$$\beta(a, b) = \sum_{r = -\infty}^{\infty} \beta_r(a, b) T^r,$$

where $\beta_r(a, b) \in C$ for every $r \in \mathbb{Z}$, and

$$\beta_r(a, b) = 0$$

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 $f(T) = \sum_{j = -\infty}^{\infty} f_j T^j \in A((T))$ and $g(T) = \sum_{l = -\infty}^{\infty} g_l T^l \in B((T))$ 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

$$\beta(f_j, g_{n-j}) = 0$$

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

$$h_n = \sum_{j = -\infty}^{\infty} \beta(f_j, g_{n-j})$$

defines an element of $C((T))$. Equivalently,

$$\beta(f_{n-l}, g_l) = 0$$
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

(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

(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

(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

(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

(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

(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

(8.6.12) \[ \beta(f(T), f(T)) = 0 \]

for every \( f(T) \in A((T)) \). To see this, it suffices to verify that

(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

$$(8.7.1) \quad \beta(f(T), g(T)) = \beta(f(T), g(T)) = \beta(f(T), g(T)) T$$

for every $f(T) \in A((T))$ and $g(T) \in B((T))$. This implies that

$$(8.7.2) \quad \beta(f(T) T^{m_1}, g(T) T^{m_2}) = \beta(f(T), g(T)) T^{m_1 + m_2}$$

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 $f(T) \in A[[T]]$ and $g(T) \in B[[T]]$ for some $m_1, m_2 \in \mathbb{Z}$, then

$$(8.7.3) \quad \beta(f(T), g(T)) \in (C[[T]]) T^{m_1 + m_2 + r(\beta)}.$$  

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

$$(8.7.4) \quad \bar{\beta}(f(T), g(T)) = \beta(f(T), g(T)) T^{-r(\beta)}$$

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 $\{a_m(T)\}_{m=m_0}$ and $\{b_r(T)\}_{r=r_0}$ 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 \(g(T) = \sum_{l=l_0}^\infty gl^i T^i\) be elements of \(A((T))\), where \(j_0, l_0 \in Z\). As usual, we take \(f_j = 0\) when \(j < j_0\), and \(g_l = 0\) when \(l < l_0\). Let \(n \in Z\) 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 \(j \in Z\), so that

\[
(8.8.2) \quad h_n = \sum_{j=-\infty}^\infty f_j g_{n-j}
\]
CHAPTER 8. FORMAL SERIES AND ORDERED RINGS

reduces to a finite sum in $A$. Equivalently,

(8.8.3) \[ f_{n-l}g_l = 0 \]

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) \[ h_n = \sum_{l=-\infty}^{\infty} f_{n-l}g_l. \]

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) \[ f(T)g(T) = h(T) = \sum_{n=j_0+l_0}^{\infty} h_n T^n, \]

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) \[ (f(T)T^{m_1})(g(T)T^{m_2}) = (f(T)g(T))T^{m_1+m_2} \]

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) \[ f(T)g(T) \in (A[[T]])T^{m_1+m_2}. \]

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
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 \(m_1, m_2 \in \mathbb{Z}\) in this situation. Similarly, if \(f(T) \in (k[[T]])^m\) and \(g(T) \in (A[[T]])^m\) for some \(m_1, m_2 \in \mathbb{Z}\), then (8.8.7) holds.

Let \(B\) be another module over \(k\), and let \(\phi(T) \in (\text{Hom}_k(A, B)((T)))\) be 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

\[
(8.8.8) \quad h_n = \sum_{j=-\infty}^{\infty} [f_j, g_{n-j}]_A
\]

for each \(n \in \mathbb{Z}\), so that (8.8.5) corresponds to

\[
(8.8.9) \quad [f(T), g(T)]_{A((T))} = h(T) = \sum_{n=-\infty}^{\infty} h_n T^n.
\]

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

\[
(8.9.1) \quad j_0(f(T) + g(T)) \geq \min(j_0(f(T)), j_0(g(T)))
\]

and

\[
(8.9.2) \quad j_0(f(T)g(T)) = j_0(f(T)) + j_0(g(T))
\]
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
\[
|f(T)|_r = r^{j_0(f(T))}
\]
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 \( r < 1 \). If \( l \in \mathbb{Z} \), 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 $x y$. 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

$$(8.10.1) \quad \phi(a) = \sum_{j=-\infty}^{\infty} \phi_j(a) T^j$$

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

$$(8.10.2) \quad \phi(a a') = \phi(a) \phi(a')$$

for every $a, a' \in A$. Let $a, a' \in A$ be given, so that

$$(8.10.3) \quad \phi(a a') = \sum_{n=-\infty}^{\infty} \phi_n(a a') T^n,$$

as in (8.10.1), where $\phi_n(a a') = 0$ for all but finitely many $n < 0$. We also have that

$$(8.10.4) \quad \phi(a) \phi(a') = \left( \sum_{j=-\infty}^{\infty} \phi_j(a) T^j \right) \left( \sum_{l=-\infty}^{\infty} \phi_l(a') T^l \right)$$

$$= \sum_{n=-\infty}^{\infty} \left( \sum_{j=-\infty}^{\infty} \phi_j(a) \phi_{n-j}(a') \right) T^n,$$

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

$$(8.10.5) \quad \phi_n(a a') = \sum_{j=-\infty}^{\infty} \phi_j(a) \phi_{n-j}(a')$$

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

$$(8.10.6) \quad \phi(f(T)) = \sum_{j=j_0}^{\infty} \phi(f_j) T^j,$$
where the sum on the right is defined as an element of $B((T))$ as in Section 8.2. Let 

$$g(T) = \sum_{l=0}^{\infty} g_l T^l$$

be another element of $A((T))$, so that

$$\phi(g(T)) = \sum_{l=0}^{\infty} \phi(g_l) T^l. \quad (8.10.7)$$

Remember that

$$h(T) = \sum_{n=j_0+l_0}^{\infty} h_n T^n$$

is defined in $A((T))$ by putting

$$h_n = \sum_{j=-\infty}^{\infty} f_j g_{n-j} \quad (8.10.8)$$

for each $n$, which reduces to a finite sum in $A$. Thus

$$\phi(h(T)) = \sum_{n=j_0+l_0}^{\infty} \phi(h_n) T^n, \quad (8.10.9)$$

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

$$\phi(h_n) = \sum_{j=-\infty}^{\infty} \phi(f_j) \phi(g_{n-j}) \quad (8.10.10)$$

for each $n$, where the sums reduce to finite sums in $B((T))$. We would like to say that

$$\sum_{n=j_0+l_0}^{\infty} \left( \sum_{j=-\infty}^{\infty} \phi(f_j) \phi(g_{n-j}) \right) T^n = \left( \sum_{j=j_0}^{\infty} \phi(f_j) T^j \right) \left( \sum_{l=l_0}^{\infty} \phi(g_l) T^l \right), \quad (8.10.11)$$

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

$$\phi(f(T) g(T)) = \phi(f(T)) \phi(g(T)), \quad (8.10.12)$$

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

$$f(T) = \sum_{j>>-\infty} f_j T^j \in A((T)),$$

then

$$\phi(f(T)) = \sum_{j>>-\infty} \phi(f_j) T^j \quad (8.11.1)$$
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

\begin{equation}
(8.11.2)
\phi(f(T) T^r) = \phi(f(T)) T^r
\end{equation}

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 $x y$. 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

\begin{equation}
(8.11.3)
f(T)^* = \sum_{j >> -\infty} f_j^* T^j
\end{equation}

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 $\mathbb{C}$. Remember that complex vector spaces may be 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 $a(T) = \sum_{j >> -\infty} a_j T^j \in \mathbb{C}((T))$, then

\begin{equation}
(8.11.4)
a(T) = \sum_{j >> -\infty} \overline{a_j} T^j
\end{equation}

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

$$x > 0.$$  

(8.12.1)

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

$$x + y > 0, \quad xy > 0.$$  

(8.12.2)

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,

$$xy = (-x)(-y) > 0$$  

(8.12.3)

when $x, y > 0$, and when $-x, -y > 0$. Similarly,

$$-xy = (-x)y = x(-y) > 0$$  

(8.12.4)

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

$$x^2 > 0,$$  

(8.12.5)

because $x$ or $-x$ is positive, and $x^2 = (-x)^2$. Of course, this corresponds to (8.12.3), with $y = x$. In particular,

$$e = e^2 > 0$$  

(8.12.6)

in $R$. If $x, y \in R$, then put

$$x < y$$  

(8.12.7)

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

$$\text{(8.13.1)} \quad n \cdot e > 0$$

in $R$, by (8.12.6). Of course, this implies that

$$\text{(8.13.2)} \quad -(n \cdot e) = (-n) \cdot e > 0$$

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

$$\text{(8.13.3)} \quad x^{-1} > 0.$$  

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

$$\text{(8.13.4)} \quad |x| = x \quad \text{when } x \geq 0 \quad \quad \quad \quad = -x \quad \text{when } -x \geq 0,$$

as on p10 of [3], and p264 of [18]. Note that $|x| \geq 0$ and

$$\text{(8.13.5)} \quad -|x| \leq x \leq |x|$$

for every $x \in R$. One can check that

$$\text{(8.13.6)} \quad |xy| = |x||y|$$

and

$$\text{(8.13.7)} \quad |x + y| \leq |x| + |y|$$

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 $\sum_{j=0}^{\infty} f_j T^j$, 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

\begin{equation}
(8.13.8) 
  h_{j_0+l_0} = f_{j_0} g_{l_0}.
\end{equation}

This implies that $h_{j_0+l_0} > 0$ in $R$, so that $h(T) > 0$ in $R((T))$.

One can verify that

\begin{equation}
(8.13.9) 
  f(T) + g(T) > 0
\end{equation}

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

\begin{equation}
(8.14.1) 
  n \cdot x > y.
\end{equation}
8.14. ORDERED FIELDS

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,

\[
n \cdot T < 1
\]

for every positive integer \( n \).

Some topics related to inner products on vector spaces over ordered fields will be discussed in Section 11.9.
Chapter 9

Solvability and nilpotency

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

\[
\psi \circ q_0 = \phi.
\]  

(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

\[
\Psi \circ q_1 = q_2,
\]  

(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
9.2. PRODUCTS OF IDEALS

Section 7.1. Let $q_4$ be the canonical quotient mapping from $A$ onto $A/A_4$. The restriction of $q_4$ 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 $q_4(A_3) = q_4(A_3 + A_4)$. The kernel of the restriction of $q_4$ to $A_3$ is equal to $A_3 \cap A_4$. The restriction of $q_4$ to $A_3$, as an algebra homomorphism from $A_3$ onto $q_4(A_3)$, 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 $q_4(A_3 + A_4)$, 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 $a_1 a_2$, 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 $[A_1, A_2]$, 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

\[ [A_1, A_2] = [A_2, A_1], \tag{9.2.5} \]

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, \(A/\lbrack A, A\rbrack\) 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 \(\lbrack a_1, a_2\rbrack_A\) 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 \(\lbrack a_1, a_2\rbrack \in A_0\), so that

\[ [A, A] \subseteq A_0. \tag{9.2.6} \]

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

\[ [B, B] \subseteq [A, A]. \tag{9.2.7} \]

### 9.3 Solvable Lie algebras

Let \(k\) be a commutative ring with a multiplicative identity element, and let \((A, \lbrack \cdot, \cdot\rbrack_A)\) 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

\[ A^{(j+1)} = [A^{(j)}, A^{(j)}] \tag{9.3.1} \]

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

\[ A^{(j+1)} \subseteq A^{(j)} \tag{9.3.2} \]

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

\[ B^{(j)} \subseteq A^{(j)} \tag{9.3.3} \]
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 \( A^{(j)} \) 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 \((A, [\cdot, \cdot], A)\) be a Lie algebra over \( k \). Also let \( A_1 \) and \( A_2 \) be ideals in \( A \), so that
$A_1 \cap A_2$ 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 $B^{(j)}$ 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 \((A, [\cdot, \cdot], A)\) 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

\[
A^{j+1} = [A, A^j]
\]

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 \( C^j A \) 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

\[
A^{j+1} \subseteq A^j
\]

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

\[
A^{(j)} \subseteq A^j
\]

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

\[
B^{j+1} = [B, B^j]
\]

for every \( j \geq 0 \). It is easy to see that

\[
B^j \subseteq A^j
\]

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

\[
\phi(A^j) = C^j
\]

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

\[
B \subseteq Z(A)
\]

if and only if

\[
[A, B] = \{0\}.
\]
If
\[(9.5.9)\quad A^j \subseteq Z(A)\]
for some \(j \geq 0\), then it follows that
\[(9.5.10)\quad A^{j+1} = [A, A^j] = \{0\},\]
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
\[(9.6.1)\quad \alpha a_0 + \beta b_0\]
9.7. NILPOTENCY CONDITIONS

for some \( \alpha, \beta \in k \). This implies that \( A^{(1)} = [A, A] \) consists of elements of \( A \) of the form

\[
\gamma [a_0, b_0]_A,
\]

where \( \gamma \in k \). It follows that \( A^{(2)} = [A^{(1)}, A^{(1)}] = \{0\} \), and in particular that \( A \) is solvable.

By hypothesis,

\[
[a_0, b_0]_A = \alpha_0 a_0 + \beta_0 b_0
\]

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

\[
[a, b]_A = a,
\]

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

\[
[a, a]_A = 0
\]

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 \( x, y, \) 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 \( x, y, \) 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 \( x, y, \) and \( z \) of \( A \) where each of \( x, y, \) and \( z \) is equal to either \( a_0 \) or \( b_0 \), because of bilinearity. This means that at least two of the elements \( x, y, \) 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

\[
\text{ad}_x(y) = [x, y]_A
\]
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 $ad_{x_j}$ is a module homomorphism from $A$ into itself for each $j = 1, \ldots, n$, so that compositions of the $ad_{x_j}$'s are defined as module homomorphisms from $A$ into itself. If $y \in A$, then

\[(9.7.2) \quad (ad_{x_1} \circ ad_{x_2} \circ \cdots \circ ad_{x_n})(y) = [x_1, [x_2, [\ldots, [x_n, y]_A \ldots]_A]_A].\]

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

\[(9.7.3) \quad ad_{x_1} \circ ad_{x_2} \circ \cdots \circ ad_{x_n} = 0\]

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 $[b, b']_B = b b' - b' b$ 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 $B^n_{\text{Lie}} = \{0\}$, where $B^n_{\text{Lie}}$ is defined as in Section 9.5. In particular, this means that $B_{\text{Lie}}$ 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 $ad_x$ 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 $(ad_x)^n = 0$ 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 $B_{\text{Lie}}$. If $b \in B$, then $\text{ad}_b$ is defined as a module homomorphism from $B$ into itself, by
\begin{equation}
\text{ad}_b(c) = [b,c]_B = bc - cb
\end{equation}
for every $c \in B$. Equivalently,
\begin{equation}
\text{ad}_b = M_b - \widetilde{M}_b,
\end{equation}
where $M_b$ and $\widetilde{M}_b$ 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 $\widetilde{M}_b$ 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 $(\widetilde{M}_b)^l = \widetilde{M}_{b^l} = 0$, 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 $B_{\text{Lie}}$ 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
\begin{equation}
\{\alpha x : \alpha \in k\}
\end{equation}
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
\begin{equation}
[x,y] \in B_0
\end{equation}
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
\begin{equation}
B_0 \subseteq N_A(B_0),
\end{equation}
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

$$\\{\alpha x + y : \alpha \in k, y \in B\}\tag{9.8.4}$$

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

$$\\{\alpha \in k : \alpha x = 0\}\tag{9.8.5}$$

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 \textit{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

$$a(v) = 0\tag{9.9.1}$$

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 $\text{ad}_x$ be the linear mapping from $gl(V)$ into itself defined by

$$\text{ad}_x(y) = [x, y]\tag{9.9.2}$$
for every $y \in gl(V)$, as in Section 2.4. If $x$ is nilpotent as an element of $L(V)$, then $ad_x$ is nilpotent as a linear mapping from $gl(V)$ into itself, as in Section 9.7. If $x \in A$, then $ad_x$ maps $A$ into itself, because $A$ is a Lie subalgebra of $gl(V)$, by hypothesis. More precisely, $ad_x$ is nilpotent as a linear mapping from $A$ into itself, because $x$ is nilpotent as an element of $L(V)$, by hypothesis.

If $x \in B$, then $ad_x$ 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

\[ T^{A/B} \circ q = q \circ T. \]

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 $T = ad_x$, to get a linear mapping $ad_x^{A/B}$ from $A/B$ into itself. More precisely, $ad_x^{A/B}$ is nilpotent on $A/B$, because $ad_x$ is nilpotent on $A$, as before. Let $C$ be the collection of $ad_x^{A/B}$, with $x \in B$. This is a Lie subalgebra of $gl(A/B)$. 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

\[ \text{ad}_x^{A/B}(q(a_0)) = 0 \]

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

\[ [x, a_0] = \text{ad}_x(a_0) \in B \]

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

\[ W = \{v \in V : x(v) = 0 \text{ for every } x \in B\}, \]

which is a linear subspace of $V$. Using the induction hypothesis again, we get that $W \neq 0$. Let us check that

\[ a(W) \subseteq W \]

for every $a \in A$. If $x \in B$ and $v \in W$, then

\[ x(a(v)) = a(x(v)) - ([a, x])(v) = 0, \]
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

\[
U_{l+1}(F) \subseteq U_l(F)
\]

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, \(U_0(F)\) is an ideal in \(U_0(F)\) for each \(l\). If \(T_1, T_2, \ldots, T_n \in U_l(F)\), 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. The proof on the middle of p13 in [13] applies the theorem mentioned in the previous section to the image of the adjoint representation of \( A \) when \( A \neq \{0\} \) to get that \( Z(A) \neq \{0\} \). One can repeat the process on \( A/Z(A) \) to get that \( A \) is nilpotent.

Let \( A \) be a finite-dimensional nipotent 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 \( gl(B) \), 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

\[
(9.11.1) \quad T_1 \circ T_2 - T_2 \circ T_1 \in \mathcal{U}_1(\mathcal{F}).
\]
Remember that $\mathcal{U}_0(\mathcal{F})$ is a Lie subalgebra of $gl(V)$, so that $\mathcal{U}_0(\mathcal{F})$ 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

$$[\mathcal{U}_0(\mathcal{F}), \mathcal{U}_0(\mathcal{F})] \subseteq \mathcal{U}_1(\mathcal{F}).$$

This implies that $\mathcal{U}_0(\mathcal{F})$ is solvable as a Lie algebra, because $\mathcal{U}_1(\mathcal{F})$ 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

$$a_{j,l} = 0$$

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

$$T_{n,r_2}(A) \subseteq T_{n,r_1}(A)$$

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

$$a b \in T_{n,r_1 + r_2}(A).$$

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

$$a b - b a \in T_{n,1}(A)$$

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

$$[t_{n,0}(A), t_{n,0}(A)] \subseteq t_{n,1}(A),$$
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 \( U_r(\mathcal{F}) \) 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

\[
9.12.1 \quad b \cdot v = \chi(b) v
\]

for every \( b \in B \). If \( a \in A \) and \( b \in B \), then \([a,b] \in B \), and in fact

\[
9.12.2 \quad \chi([a,b]) = 0.
\]

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

\[
9.12.3 \quad ([a,b]) \cdot v = \rho_{[a,b]}(v) = 0
\]

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

\[
9.12.4 \quad v, \rho_a(v), \ldots, \rho_a^{j-1}(v)
\]

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

\[
9.12.5 \quad V_n = V_{n+1}.
\]

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

\[
9.12.6 \quad \rho_a(V_n) \subseteq V_{n+1} = V_n.
\]
In particular, $n$ is less than or equal to the dimension of $V$.

If $b \in B$, then we would like to show that

$\rho_b((\rho_a)^j(v)) = \chi(b) (\rho_a)^j(v) \mod V_j$

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

$\rho_b((\rho_a)^j(v)) = \rho_b(\rho_a((\rho_a)^{j-1}(v)))$

$= \rho_a(\rho_b((\rho_a)^{j-1}(v))) - ([\rho_a, \rho_b])(\rho_a)^{j-1}(v))$

$= \rho_a(\rho_b((\rho_a)^{j-1}(v))) - \rho_{[a,b]}((\rho_a)^{j-1}(v)).$

Of course,

$\rho_b((\rho_a)^{j-1}(v)) = \chi(b) (\rho_a)^{j-1}(v) \mod V_{j-1},$

by the induction hypothesis. This implies that

$\rho_a(\rho_b((\rho_a)^{j-1}(v))) = \chi(b) (\rho_a)^j(v) \mod V_j,$

because $\rho_a(V_{j-1}) \subseteq V_j$ by construction. Similarly,

$\rho_{[a,b]}((\rho_a)^{j-1}(v)) = \chi([a,b]) (\rho_a)^j(v) \mod V_{j-1},$

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

$\text{tr}_{V_n} \rho_b = n \cdot \chi(b)$

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

$\text{tr}_{V_n} \rho_{[a,b]} = \text{tr}_{V_n} ([\rho_a, \rho_b]) = 0$

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 an $v \in V$ 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 an $v \in V$ with $v \neq 0$ and an $a \in A$ such that

$$\rho_a(W) \subseteq W$$

for every $a \in A$. If $a \in A$, $b \in B$, and $w \in W$, then

$$\rho_a(\rho_b(w)) = \rho_a(\chi(b) w) = \chi(b) \rho_a(w) - \chi([a, b]) w,$$

using the fact that $[a, b] \in B$ in the second step. Combining this with (9.12.2), we obtain that

$$\rho_b(\rho_a(w)) = \chi(b) \rho_a(w),$$

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 $\rho_{a_0}$, as desired, because $w_0$ is an eigenvector for $\rho_b$, 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 $\mathcal{F} = \{V_j\}_{j=0}^n$ 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 $u_i$ 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

$$[u_j, u_l]_k^n = \sum_{r=1}^n c^r_{j,l} u_r$$  

(9.14.2)

for each $j, l = 1, \ldots, n$, where $c^r_{j,l}$ are elements of $k$ for every $j, l, r = 1, \ldots, n$. This implies that

$$([x, y]_k^n)_r = \sum_{j=1}^n \sum_{l=1}^n c^r_{j,l} x_j y_l$$  

(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 $c^r_{j,l} \in k$ for each $j, l, r = 1, \ldots, n$, 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

$$c^r_{j,l} = -c^r_{l,j}$$  

(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

\[
(9.14.7) \quad c^r_{j,j} = 0
\]

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

\[
(9.14.8) \quad \sum_{h=1}^{n} (c^h_{j,l} c^h_{h,m} + c^h_{l,m} c^h_{h,j} + c^h_{m,j} c^h_{h,l}) = 0
\]

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 \( c^r_{j,l} \)'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

\[
(9.14.9) \quad ([a, b]_{A^n})_r = \sum_{j=1}^{n} \sum_{l=1}^{n} c^r_{j,l} a_j b_l
\]

for each \( r = 1, \ldots, n \), where the left side is the \( r \)th coordinate of \([a, b]_{A^n}\). The conditions on the \( c^r_{j,l} \)'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,

\[
(9.14.10) \quad t \mapsto te
\]

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

\[
(9.15.1) \quad (\text{ad}_A x)(y) = \text{ad}_{A,x}(y) = [x, y]
\]
from $A$ into itself. Similarly, if $B$ is a Lie subalgebra of $A$ and $x \in B$, then $\text{ad}_{B,x}$ 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 nipotent 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 nipotent 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 $[k^n, k^n]$ is nipotent 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 $gl_n(k)$

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

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

(10.1.1)

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

$$
x = \sum_{r=1}^{n} x_r u_r
$$

(10.1.2)

for every $x = (x_1, \ldots, x_n) \in k^n$. If $a \in M_n(k)$, then

$$
(T_a(u_r))_j = a_{j,r}
$$

(10.1.3)

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

$$
a = \sum_{h=1}^{n} \sum_{m=1}^{n} a_{h,m} e_{h,m}.
$$

(10.1.4)
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

$$T_{h,m} = T_{e_{h,m}}. \quad (10.1.5)$$

If $a = (a_{j,l}) \in M_n(k)$, then $T_a$ can be expressed as

$$T_a = \sum_{h=1}^{n} \sum_{m=1}^{n} a_{h,m} T_{h,m}. \quad (10.1.6)$$

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

$$T_{h,m}(u_r) = \delta_{m,r} u_h \quad (10.1.7)$$

for every $h, m, r = 1, \ldots, n$. We also have that

$$e_{h,m} e_{q,r} = \delta_{m,q} e_{h,r} \quad (10.1.8)$$

for every $h, m, q, r = 1, \ldots, n$. Equivalently,

$$T_{h,m} \circ T_{q,r} = \delta_{m,q} T_{h,r} \quad (10.1.9)$$

for every $h, m, q, r = 1, \ldots, n$. It follows from (10.1.8) that

$$[e_{h,m}, e_{q,r}] = e_{h,m} e_{q,r} - e_{q,r} e_{h,m} = \delta_{m,q} e_{h,r} - \delta_{r,h} e_{q,m} \quad (10.1.10)$$

for every $h, m, q, r = 1, \ldots, n$. In particular, if $h \neq r$ and $m \neq q$, then

$$[e_{h,m}, e_{q,r}] = 0, \quad (10.1.11)$$

because each of the two terms on the right side of (10.1.10) is equal to 0. Otherwise,

$$[e_{h,m}, e_{m,r}] = e_{h,r} \quad (10.1.12)$$

when $h \neq r$, and

$$[e_{h,m}, e_{q,h}] = -e_{q,m} \quad (10.1.13)$$

when $m \neq q$. Of course, these two cases are equivalent, because of the antisymmetry of the commutator bracket. Similarly,

$$[e_{h,m}, e_{m,h}] = e_{h,h} - e_{m,m} \quad (10.1.14)$$

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,

$$[gl_n(k), gl_n(k)] = sl_n(k), \quad (10.1.15)$$
10.2. SOME BASIC PROPERTIES OF $SL_2(K)$

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

$$(10.2.1) \quad x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$ 

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

$$(10.2.2) \quad x = e_{1,2}, \quad y = e_{2,1}, \quad \text{and} \quad h = e_{1,1} - e_{2,2}.$$ 

One can check that

$$(10.2.3) \quad [x, y] = h, \quad [h, x] = 2 \cdot x, \quad [h, y] = -2 \cdot y,$$

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

$$[h, x] = [h, y] = 0,$$

(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

$$(10.2.5) \quad [sl_2(k), sl_2(k)] = sl_2(k),$$

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

\[(10.2.6) \quad [sl_n(k), sl_n(k)] = sl_n(k) . \]

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

\[(10.2.7) \quad e_{h,r} \in [sl_n(k), sl_n(k)] \]

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

\[(10.2.8) \quad e_{h,h} - e_{m,m} \in [sl_n(k), sl_n(k)] \]

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 $a = (a_{i,j})$ with entries in $A$ is said to be a diagonal matrix if $a_{j,i} = 0$ when $j \neq i$. 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

\[(10.3.1) \quad e_{h,m} \in sl_n(k) \]

for every $h, m = 1, \ldots, n$ with $h \neq m$, and

\[(10.3.2) \quad e_{h,h} - e_{m,m} \in sl_n(k) \]
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 $sl_n(k)$. One can check that every element of $sl_n(k)$ 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 $sl_n(k)$. Thus $sl_n(k)$ 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 $gl_n(k)$ 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 $gl_n(k)$ is isomorphic to the free module $k^{n^2}$ of rank $n^2$ over $k$, as a module over $k$. Alternatively, one can verify that every element of $gl_n(k)$ can be expressed in a unique way as a linear combination of the standard basis elements for $sl_n(k)$ mentioned in the preceding paragraph together with $e_{1,1}$, with coefficients in $k$. In particular, every element of $gl_n(k)$ can be expressed in a unique way as the sum of an element of $sl_n(k)$ and a multiple of $e_{1,1}$ by an element of $k$, so that $gl_n(k)$ is isomorphic to the direct sum of $sl_n(k)$ and $k$, as modules over $k$.

Indeed, if $a \in gl_n(k)$, then
\[
(10.3.3) \quad a - (\text{tr } a) e_{1,1} \in sl_n(k),
\]
because $\text{tr } e_{1,1} = 1$. Thus $a$ can be expressed as the sum of (10.3.3) and $(\text{tr } a) e_{1,1}$, and one can use the trace again to see that this is unique.

Note that $\text{tr } I = n - 1$, 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 $a \in gl_n(k)$, then
\[
(10.3.4) \quad a - \frac{\text{tr } a}{n \cdot 1} I \in sl_n(k),
\]
and $\text{tr } a/(n \cdot 1)$ is the unique element of $k$ with this property. This implies that every element of $gl_n(k)$ can be expressed in a unique way as the sum of elements of $sl_n(k)$ and $s_n(k)$, as in Exercise 7 on p5 in [13].

## 10.4 Centrality in $gl_n(k)$, $sl_n(k)$

Let $k$ be a commutative ring with a multiplicative identity element, and let $n$ be a positive integer. Remember that $a = (a_{j,l}) \in gl_n(k)$ can be expressed as
\[
(10.4.1) \quad a = \sum_{h=1}^{n} \sum_{m=1}^{n} a_{h,m} e_{h,m},
\]
as in (10.1.4). If $q, r \in \{1, \ldots, n\}$, then
\[
[a, e_{q,r}] = \sum_{h=1}^{n} \sum_{m=1}^{n} a_{h,m} [e_{h,m}, e_{q,r}]
\]
\[
(10.4.2) = \sum_{h \neq r} a_{h,q} e_{h,r} - \sum_{m \neq q} a_{r,m} e_{q,m} + a_{r,q} (e_{r,r} - e_{q,q}),
\]
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

\[
[a, e_{q,r}] = \sum_{h \neq q,r} a_{h,q} e_{h,r} - \sum_{m \neq q,r} a_{r,m} e_{q,m} + (a_{q,q} - a_{r,r}) e_{q,r}
\]

(10.4.3)

+ \( a_{r,q} (e_{r,r} - e_{q,q}) \),

where the sums are taken over \( h, m = 1, \ldots, n \) with \( h, m \neq q, r \), 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 \( t I \), where \( t \in k \) satisfies \( n \cdot t = 0 \). If \( n \cdot 1 = 0 \) 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

\[
[a, e_{r,r}] = \sum_{h \neq r} a_{h,r} e_{h,r} - \sum_{m \neq r} a_{r,m} e_{r,m},
\]

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

\[
\text{tr}(T \circ R) = 0,
\]

(10.5.1)
as in Theorem 7.1 on p.42 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 \(\mathcal{F} = \{V_j\}_{j=0}^n\) 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 \(\mathcal{F}\). This corresponds to Exercise 7 on p.21 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 \(gl_n(k_1)\) 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

$$T(R(v)) = R(T(v)) = \lambda R(v).$$

This means that $R(v) \in E_\lambda(T)$, so that

$$R(E_\lambda(T)) \subseteq E_\lambda(T).$$

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

$$\prod_{j \neq r} (T - \lambda_j I)$$

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 $E_\lambda_j(T)$ when $j \neq r$, and it is equal to

$$\prod_{j \neq r} (\lambda_r - \lambda_j)$$

times the identity mapping on $E_\lambda_r(T)$. As before, (10.6.5) is the product of $\lambda_r - \lambda_j$ for $j = 1, \ldots, l$ and $j \neq r$, which is a nonzero element of $k$. If $v_j \in E_\lambda_j(T)$ for each $j = 1, \ldots, l$ and

$$\sum_{j=1}^l v_j = 0,$$

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

$$T(W) \subseteq W.$$ 

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.
10.7. NILPOTENT VECTORS

Let $R$ be another linear mapping from $V$ into itself that commutes with $T$ again. Thus $R$ maps $E_{\lambda}(T)$ 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 $E_{\lambda}(T)$ 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 $1$, where multiplication of $a, b \in A$ is expressed as $a b$. 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

\[(I - T) \sum_{j=0}^{n} T^j = \left( \sum_{j=0}^{n} T^j \right) (I - T) = I - T^{n+1} \tag{10.7.1}\]

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

\[\sum_{j=0}^{\infty} T^j(v)\]

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

\[(I - T) \sum_{j=0}^{\infty} T^j = \left( \sum_{j=0}^{\infty} T^j \right) (I - T) = I, \tag{10.7.3}\]

using (10.7.1). This implies that $I - T$ is invertible on $V$, with

\[(I - T)^{-1} = \sum_{j=0}^{\infty} T^j. \tag{10.7.4}\]

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

\[E_0(T) = \{ v \in V : T^l(v) = 0 \text{ for some } l \in \mathbb{Z}_+ \}. \tag{10.7.5}\]

This is a linear subspace of $V$ that contains the kernel of $T$. If $v \in E_0(T)$ and $v \neq 0$, and if $j$ is the smallest positive integer such that $T^j(v) = 0$, then $T^{j-1}(v)$
is a nonzero element of the kernel of $T$. Thus $E_0(T) \neq \{0\}$ if and only if the kernel of $T$ is nontrivial. Note that $T$ maps $E_0(T)$ into itself. If $\lambda \in k$ and $\lambda \neq 0$, then $\lambda I - T$ is invertible as a linear mapping on $E_0(T)$, by the remarks in the previous paragraph. If $R$ is a linear mapping from $V$ into itself that commutes with $T$ and $v \in E_0(T)$, then

$$T^l(R(v)) = R(T^l(v)) = 0$$

for some $l \in \mathbb{Z}_+$. This means that $R(v) \in E_0(T)$, and hence

$$R(E_0(T)) \subseteq E_0(T).$$

Let $\lambda \in k$ be given, and put

$$E_\lambda(T) = E_0(T - \lambda I) = \{v \in V : (T - \lambda I)^l(v) = 0 \text{ for some } l \in \mathbb{Z}_+\}.$$ 

This is a linear subspace of $V$, which reduces to (10.7.5) when $\lambda = 0$. Of course,

$$E_\lambda(T) \subseteq E_\lambda(T),$$

where $E_\lambda(T)$ is as in (10.6.1). As before, $E_\lambda(T) \neq \{0\}$ if and only if the kernel of $T - \lambda I$ is nontrivial, which means that $\lambda$ is an eigenvalue of $T$. If $R$ is a linear mapping from $V$ into itself that commutes with $T$, then

$$R(E_\lambda(T)) \subseteq E_\lambda(T),$$

as in (10.7.7). In particular, $T$ maps $E_\lambda(T)$ 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 $E_\lambda(T)$ 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 $v_j \in E_{\lambda_j}(T)$ 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 $E_{\lambda_j}(T)$ 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

$$\text{(10.8.2)} \quad E_\lambda(T_1) \subseteq E_\lambda(T),$$

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

$$\text{(10.8.3)} \quad E_{\lambda_j}(T_1) = E_{\lambda_j}(T)$$

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

$$\text{(10.8.4)} \quad T = T_3 + T_4,$$

$T_3$ 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,

$$\text{(10.8.5)} \quad T_1 - T_3 = T_4 - T_2,$$

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.
Suppose that \( W_0 \) and \( W \) are linear subspaces of \( V \) with
\[
T(W) \subseteq W_0 \subseteq W. \tag{10.8.6}
\]
Under these conditions,
\[
T_1(W), T_2(W) \subseteq W_0, \tag{10.8.7}
\]
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
\[
ad_A(B) = [A, B] = A \circ B - B \circ A, \tag{10.9.1}
\]
as before. This defines \( ad_A \) as a linear mapping from \( \mathcal{L}(V) \) into itself, where \( \mathcal{L}(V) \) is the space of linear mappings from \( V \) into itself. If \( A \) is diagonalizable on \( V \), then \( ad_A \) is diagonalizable on \( \mathcal{L}(V) \). 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 \( ad_A \) is nilpotent on \( \mathcal{L}(V) \), as in Section 9.7. Let \( T, T_1, \) and \( T_2 \) be as in (10.8.1), so that
\[
ad_T = ad_{T_1} + ad_{T_2}. \tag{10.9.2}
\]
Note that \( ad_{T_1} \) is diagonalizable on \( \mathcal{L}(V) \), and that \( ad_{T_2} \) is nilpotent on \( \mathcal{L}(V) \), by the corresponding properties of \( T_1 \) and \( T_2 \) on \( V \), and the previous remarks. We also have that
\[
[ad_{T_1}, ad_{T_2}] = ad_{[T_1, T_2]} = 0, \tag{10.9.3}
\]
so that \( ad_{T_1} \) and \( ad_{T_2} \) commute on \( \mathcal{L}(V) \), because \( T_1 \) and \( T_2 \) commute on \( V \). Thus (10.9.2) is the analogue of (10.8.1) for \( ad_T \), 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
\[
\delta(ab) = \delta(a)b + a\delta(b) = (\lambda a)b + a(\mu b) = (\lambda + \mu)(ab), \tag{10.9.4}
\]
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 6 on p12 of [13] for derivations on Lie algebras that come from the adjoint representation.

If \( \lambda \in k \), then put
\[
\mathcal{E}_\lambda(\delta) = \{ a \in A : (\delta - \lambda I)^l(a) = 0 \text{ for some } l \in \mathbb{Z}_+ \}, \tag{10.9.5}
\]
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. If $a \in \mathcal{E}_\lambda(\delta)$ and $b \in \mathcal{E}_\mu(\delta)$ for some $\lambda, \mu \in k$, then

\[(10.9.6) \quad (\delta - (\lambda + \mu) I)(ab) = \delta(a) b + a \delta(b) - \lambda a b - \mu a b = ((\delta - \lambda I)(a)) b + a ((\delta - \mu I)(b)).\]

One can use this repeatedly to get that

\[(10.9.7) \quad (\delta - (\lambda + \mu) I)^j(ab) = 0\]

when $j$ is sufficiently large, so that $ab \in \mathcal{E}_{\lambda+\mu}(\delta)$. If $ab \neq 0$, 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,

\[(10.9.9) \quad \mathcal{E}_\lambda(\delta_1) = \mathcal{E}_\lambda(\delta)\]

for every $\lambda \in k$, and in particular the eigenvalues of $\delta$ and $\delta_1$ are the same. If $a \in \mathcal{E}_\lambda(\delta_1) = \mathcal{E}_\lambda(\delta)$ and $b \in \mathcal{E}_\mu(\delta_1) = \mathcal{E}_\mu(\delta)$ for some $\lambda, \mu \in k$, then $ab$ is an element of $\mathcal{E}_{\lambda+\mu}(\delta) = \mathcal{E}_{\lambda+\mu}(\delta_1)$, as in the preceding paragraph. Note that $ab$ may be equal to 0, in which case $\lambda + \mu$ need not be an eigenvalue of $\delta, \delta_1$. In this situation,

\[(10.9.10) \quad \delta_1(ab) = (\lambda + \mu) a b = (\lambda a)b + a(\mu b) = \delta_1(a)b + a \delta_1(b).\]

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 $\mathcal{E}_\lambda(A)$ 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
\[
A(v_j) = a_j v_j
\]
for every \( j = 1, \ldots, n \). Using this basis, \( \phi(A) \) can be characterized by
\[
(\phi(A))(v_j) = \phi(a_j) v_j
\]
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
\[
E_{h,m}(v_j) = \delta_{m,j} v_h
\]
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
\[
\text{ad}_A(E_{h,m}) = (a_h - a_m) E_{h,m}
\]
for every \( h, m = 1, \ldots, n \), as in (10.4.5). Similarly,
\[
\text{ad}_{\phi(A)}(E_{h,m}) = (\phi(a_h) - \phi(a_m)) E_{h,m}
\]
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
\[
\text{ad}_{\phi(A)}(E_{h,m}) = \phi(a_h - a_m) E_{h,m}
\]
for every \( h, m = 1, \ldots, n \). This means that
\[
\text{ad}_{\phi(A)} = \phi(\text{ad}_A),
\]
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 $\mathbb{Q}$. The corresponding linear mappings $\phi(A)$ are called replicas of $A$ by Chevalley, as mentioned on p42 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

\begin{equation}
T = T_1 + T_2,
\end{equation}

$T_1$ 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

\begin{equation}
\text{tr}(T \circ \phi(T_1)) = 0
\end{equation}

for all such $\phi$, then $T_1 = 0$, so that $T = T_2$ is nilpotent on $V$. This corresponds to Lemma 6.7 on p42 of [24], rephrased a bit in the way that it is used in the proof of the lemma on p19 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

\begin{equation}
\text{tr}(T_2 \circ \phi(T_1)) = 0.
\end{equation}

Thus (10.11.2) is the same as saying that

\begin{equation}
\text{tr}(T_1 \circ \phi(T_1)) = 0.
\end{equation}

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

\begin{equation}
\sum_{j=1}^{l} m_j \cdot \lambda_j \phi(\lambda_j) = 0,
\end{equation}

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

\begin{equation}
\sum_{j=1}^{l} m_j \cdot \phi(\lambda_j)^2 = 0.
\end{equation}
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 \( Q \). 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 \( Q \) into the subfield of \( k \) corresponding to \( Q \) that are linear over \( Q \). Thus \( T_1 = 0 \) on \( V \), because \( T_1 \) is diagonalizable on \( V \), by hypothesis, as desired.

If \( k \) is the field \( C \) 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 \( C \), then

\[
\sum_{j=1}^{t} m_j |\lambda_j|^2 = 0.
\]

This implies directly that 0 is the only eigenvalue of \( T_1 \), as desired.

Let \( A \) and \( B \) be linear subspaces of the space \( L(V) \) of all linear mappings from \( V \) into itself, with

\[
A \subseteq B.
\]

Put

\[
M = \{ T \in L(V) : [T, B] \in A \text{ for every } B \in B \}.
\]

If \( T \in M \) satisfies

\[
\text{tr}(T \circ R) = 0
\]

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 \( T_1, T_2 \in L(V) \) 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 \( Q \) into \( k \) that is linear over \( Q \), 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

\[
\phi(T_1) \in M.
\]

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 \( L(V) \) into itself, as in (10.9.1). The condition that \( T \in M \) means exactly that

\[
\text{ad}_T(B) \subseteq A.
\]

Remember that

\[
\text{ad}_T = \text{ad}_{T_1} + \text{ad}_{T_2}
\]

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

\[
\text{ad}_{T_1}(B) \subseteq A,
\]
so that $T_1 \in \mathcal{M}$.

We also have that
\begin{equation}
\text{ad}_{\phi(T_1)} = \phi(\text{ad}_{T_1}),
\end{equation}
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
\begin{equation}
\text{ad}_{\phi(T_1)}(\mathcal{B}) \subseteq A,
\end{equation}
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 $\text{gl}(V)$, and suppose that
\begin{equation}
\text{tr}(T \circ R) = 0
\end{equation}
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 Section 9.10.

Put
\begin{equation}
\mathcal{M} = \{ Z \in \text{gl}(V) : [Z, T] \in [A, A] \text{ for every } T \in A \},
\end{equation}
and observe that $A \subseteq \mathcal{M}$. We would like to show that
\begin{equation}
\text{tr}(R \circ Z) = 0
\end{equation}
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
\begin{equation}
\text{tr}([T_1, T_2] \circ Z) = 0
\end{equation}
for every $T_1, T_2 \in A$ and $Z \in \mathcal{M}$. Remember that
\begin{equation}
\text{tr}([T_1, T_2] \circ Z) = -\text{tr}(T_2 \circ ([T_1, Z])),
\end{equation}
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
can be reformulated in terms of matrices as saying that if $A_0$ is a Lie subalgebra of $\text{gl}_n(k)$ such that
\begin{equation}
\text{tr}(T_0 R_0) = 0
\tag{10.12.6}
\end{equation}
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 $\text{gl}_n(k)$ may be considered as a subset of $\text{gl}_n(k_1)$. If $A_1$ is the linear span of $A_0$ in $\text{gl}_n(k_1)$, as a vector space over $k_1$, then $A_1$ is a Lie subalgebra of $\text{gl}_n(k_1)$, as 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
\begin{equation}
\text{tr}(T_1 R_1) = 0
\tag{10.12.7}
\end{equation}
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
\begin{equation}
\text{ad}_x(z) = \text{ad}_{A,x}(z) = [x, z]
\tag{10.13.1}
\end{equation}
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
\begin{equation}
\beta(x, y) = \text{tr}_A(\text{ad}_x \circ \text{ad}_y)
\tag{10.13.2}
\end{equation}
is defined as an element of $k$ for every $x, y \in A$. Remember that the radical
\begin{equation}
A^2 = \{x \in A : \beta(x, y) = 0 \text{ for every } y \in A\}
\tag{10.13.3}
\end{equation}
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
\begin{equation}
\{\text{ad}_x : x \in A_0\}
\tag{10.13.4}
\end{equation}
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
10.14. COMPLEMENTARY IDEALS

Let $k$ be a commutative ring with a multiplicative identity element, and let $(A, ['\cdot', '\cdot'])$ be a Lie algebra over $k$. Suppose that $A$ is isomorphic to $k^r$ 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

$$\beta([w, x], y) = \beta(x, [w, y])$$

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

$$B^\perp = B^{\perp, \beta} = \{x \in A : \beta(x, y) = 0 \text{ for every } y \in B\}.$$  

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

\[(10.14.3) \quad A_0 = B \cap B^\perp \]

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

\[(10.14.4) \quad [x, y] = 0 \quad \text{for every } x \in B \text{ and } y \in B^\perp, \]

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

\[(10.14.5) \quad B + B^\perp = \{x + y : x \in B, y \in B^\perp\} = A. \]

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

$$[a_j, a_l] = 0,$$

(10.15.1)

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 $k$ 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 p67 of [13]. More precisely, suppose that $A$ is a nonzero ideal in $sl_2(k)$. Let

$$a x + b y + c h$$

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
11.2. TRACES ON $M_n(K)$

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

$$L_a(x) = a x$$

(11.2.1)

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

$$L_a \circ L_b = L_{a b},$$

(11.2.2)

as in Section 2.8. It is not difficult to see that

$$\text{tr}_{M_n(k)} L_a = n \cdot \text{tr} a$$

(11.2.3)

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

$$\text{tr}_{M_n(k)}(L_a \circ L_b) = \text{tr}_{M_n(k)} L_{a b} = n \cdot \text{tr}(a b)$$

(11.2.4)

for every $a, b \in M_n(k)$, using (11.2.2) in the first step.

Similarly, if $a \in M_n(k)$, then

$$R_a(x) = x a$$

(11.2.5)

defines $R_a$ as a module homomorphism from $M_n(k)$ into itself. If $b \in M_n(k)$ too, then

$$R_a \circ R_b = R_{b a},$$

(11.2.6)
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

\[(11.2.8) \quad \text{tr}_{M_n(k)} (R_a \circ R_b) = \text{tr}_{M_n(k)} (R_{ba}) = n \cdot \text{tr}(b \cdot a) = n \cdot \text{tr}(a \cdot b)\]

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

\[(11.2.10) \quad \text{ad}_a(x) = [a, x] = a x - x a = L_a(x) - R_a(x)\]

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

\[(11.2.12) \quad \text{ad}_a \circ \text{ad}_b = (L_a - R_a) \circ (L_b - R_b) = L_a \circ L_b - L_a \circ R_b - R_b \circ L_a + R_a \circ R_b = L_{a \cdot b} - L_a \circ R_b - L_b \circ R_a + R_a \circ R_b.\]

It follows that

\[(11.2.13) \quad \text{tr}_{M_n(k)} (\text{ad}_a \circ \text{ad}_b) = \text{tr}_{M_n(k)} L_{a \cdot b} - \text{tr}_{M_n(k)} (L_a \circ R_b) - \text{tr}_{M_n(k)} (L_b \circ R_a) + \text{tr}_{M_n(k)} R_{ba}\]

\[= n \cdot \text{tr}(a \cdot b) - 2 \cdot (\text{tr} a) (\text{tr} b) + n \cdot \text{tr}(b \cdot a) = 2 n \cdot \text{tr}(a \cdot b) - 2 \cdot (\text{tr} a) (\text{tr} b).\]

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

\[(11.2.15) \quad \text{tr}_{sl_n(k)} (\text{ad}_{sl_n(k),a} \circ \text{ad}_{sl_n(k),b}) = 2 n \cdot \text{tr}(a \cdot b),\]

by (11.2.13).
11.3 A nice criterion

Let \( k \) be a commutative ring with a multiplicative identity element, and let \((A, [\cdot, \cdot], \lambda)\) 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 equal to \( Z(A) \), 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 finite dimension. Also let \( A \) be a Lie subalgebra 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 \( I = I_V \) on \( V \). In particular, this holds when \( A \) is commutative or 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 \( v_0 \in V \) such that \( v_0 \neq 0 \) and \( v_0 \) 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

\[(11.3.1) \quad T(v_0) = \lambda(T) v_0 \]

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,

\[(11.3.2) \quad \lambda([R, T]) = 0, \]

as in Section 9.12.

Put

\[(11.3.3) \quad V_\lambda = \{v \in V : T(v) = \lambda(T) v \quad \text{for every} \quad T \in \text{Rad} \ A\}. \]

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

\[(11.3.4) \quad T(R(v)) = R(T(v)) - ([R, T])(v) = R(\lambda(T) v) - \lambda([R, T]) v = \lambda(T) R(v). \]

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 \( c_{j,l}^r \) be an element of \( k \) for every \( j, l, r = 1, \ldots, n \). If \( x, y \in k^n \), then let \([x, y]_k\) be the element of \( k^n \) whose \( r \)th coordinate is given by

\[
([x, y]_k)_r = \sum_{j=1}^{n} \sum_{l=1}^{n} c_{j,l}^r x_j y_l
\]

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

\[
c_{j,l}^r = -c_{l,j}^r \quad \text{and} \quad c_{j,j}^r = 0
\]

for every \( j, l, r = 1, \ldots, n \), so that \([x, x]_k = 0\) for every \( x \in k^n \). Suppose also that the \( c_{j,l}^r \)'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 \( k^m \times \{0\} \) to denote the space of \( x \in k^n \) such that \( x_j = 0 \) when \( j \geq m + 1 \). Suppose that for every \( r = 1, \ldots, n \),

\[
c_{j,l}^r = 0 \quad \text{when} \quad j \geq m + 1 \quad \text{or} \quad l \geq m + 1.
\]

This implies that \( k^m \times \{0\} \) 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 \( k^m \times \{0\} \) 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 \([a, a']_K\) be the element of \( K^n \) whose \( r \)th coordinate is equal to

\[
([a, a']_K)_r = \sum_{j=1}^{n} \sum_{l=1}^{n} c_{j,l}^r a_j a'_l
\]

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 \( k^m \times \{0\} \) is an ideal in \( K^n \) with respect to (11.4.4), because of (11.4.3). If \( k^m \times \{0\} \) is solvable as a Lie algebra over \( k \), then \( K^m \times \{0\} \) 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 $c_{r,j,l} \in k$ 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 $c_{r,j,l}$ 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 $k^m \times \{0\}$ in $k^n$. If $A_0$ is solvable as a Lie algebra over $k$, then $k^m \times \{0\}$ is solvable as a Lie algebra over $k$, and $K^m \times \{0\}$ 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

\[
\mu = \sum_{l=1}^{n} \mu(v_l) \lambda_l.
\]

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

\[
\beta_{v_l} = \sum_{j=1}^{n} \beta_{v_l}(v_j) \lambda_j = \sum_{j=1}^{n} \beta(v_j, v_l) \lambda_j
\]

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 $c_{r,j,l} \in k$ for $j, l, r = 1, \ldots, n$ that satisfy (11.4.2) and (9.14.8) such that (11.4.1)
corresponds to $[,]_A$ 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 (11.5.6) 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 $u_i$ 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}
11.6. BILINEAR FORMS AND ADJUNTS

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 $(\beta_{j,l})$ 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 $(\beta_{j,l})$ 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 $(\beta_{j,l})$.

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 $(\beta_{j,l})$ 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

$$a^* = (\beta^{-1} a \beta)^t = \beta^t a^t (\beta^t)^{-1}$$

(11.6.2)

for each $a \in M_n(k)$. It is easy to see that $a \mapsto a^*$ defines an opposite algebra automorphism on $M_n(k)$, because of the same property of $a \mapsto a^t$.

If $a = (a_{j,l}) \in M_n(k)$, then

$$\langle T_a(x) \rangle_j = \sum_{l=1}^n a_{j,l} x_l$$

(11.6.3)

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,

$$\beta(T_a(x), y) = \beta(x, T_{a^*}(y))$$

(11.6.4)

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

$$\langle (a^*)^* \rangle_j = \beta^t (a^*)^t (\beta^t)^{-1} = \beta^t (\beta^t)^{-1} (a^t)^t (\beta^t)^{-1} = \beta^t \beta^{-1} a \beta (\beta^t)^{-1}$$

(11.6.5)

for every $a \in M_n(k)$. If $\beta$ is either symmetric or antisymmetric, then we get that

$$\langle (a^*)^* \rangle_j = a$$

(11.6.6)

for every $a \in M_n(k)$, and (11.6.2) is the same as

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

(11.6.7)

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 $\beta = (\beta_{j,l})$ 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

$$\det a = \det a^t = \det(-a) = (-1)^n \det a. \tag{11.7.1}$$

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 $\beta(\cdot, T_a(\cdot))$ 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

$$\beta(T_a(x), y) = \beta(x, T_a(y)) \tag{11.7.3}$$

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

$$\beta(T_a(x), y) = \beta(T_a(y), x) \tag{11.7.4}$$

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

$$\beta(T_a(x), y) = -\beta(T_a(y), x) \tag{11.7.5}$$

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

$$\beta(T_a(x), y) = -\beta(x, T_a(y)) \tag{11.7.6}$$

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 $\gamma = (\gamma_{j,l}) \in M_n(k)$, 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 $k^n$.

(11.7.7) \[
\gamma(x, y) = \sum_{j=1}^{n} \sum_{l=1}^{n} \gamma_{j,l} x_j y_l
\]
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) \[
\gamma = \beta a.
\]
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) \[
\text{tr } a^* = \text{tr } a
\]
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) \[
(a, b)_{M_n(k)} = \text{tr}(ab^*)
\]
for every $a, b \in M_n(k)$, which defines a bilinear form on $M_n(k)$. Observe that

(11.8.3) \[
(a, b)_{M_n(k)} = \text{tr}(ab^*) = \text{tr}((b^*)^* a^*) = \text{tr}(b a^*) = (b, a)_{M_n(k)}
\]
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) \[
L_c(a) = ca \quad \text{and} \quad R_c(a) = ac
\]
be the corresponding operators of left and right multiplication on $M_n(k)$ by $c$, respectively. Thus

(11.8.5) \[
(L_c(a), b)_{M_n(k)} = \text{tr}(ca b^*) = \text{tr}(a b^* c)
\]
for every $a, b \in M_n(k)$. Similarly,

(11.8.6) \[
(R_c(a), b)_{M_n(k)} = \text{tr}(a c b^*) = \text{tr}(a (b c^*)^*) = (a, R_c(b))_{M_n(k)}
\]
for every $a, b \in M_n(k)$. Let $a$ and $b$ 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

\[(a, b)_{M_n(k)} = \text{tr}(a b). \tag{11.8.7}\]

If $a$ is self-adjoint and $b$ is arbitrary, then

\[(a, b)_{M_n(k)} = (b, a)_{M_n(k)} = \text{tr}(b a) = \text{tr}(a b). \tag{11.8.8}\]

If $a$ is arbitrary and $b$ is anti-self-adjoint, so that $b^* = -b$, then

\[(a, b)_{M_n(k)} = -\text{tr}(a b). \tag{11.8.9}\]

If $c \in M_n(k)$, then put

\[C_c(a) = [c, a] = c a - a c = L_c(a) - R_c(a) \tag{11.8.11}\]

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

\[(C_c(a), b)_{M_n(k)} = (a, C_c^*(b))_{M_n(k)} \tag{11.8.12}\]

for every $a, b \in M_n(k)$, by (11.8.5) and (11.8.6). Of course,

\[(C_c(a))^* = ([c, a])^* = -[c^*, a^*] = -C_c^*(a^*) \tag{11.8.13}\]

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

\[(a, b)_{M_n(k)} = 0, \tag{11.8.14}\]

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

\[\langle v, v \rangle_V > 0 \tag{11.9.1}\]

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

\[\langle v, w \rangle_V = 0. \tag{11.9.2}\]
Suppose that $e_1, \ldots, e_n$ are finitely many nonzero vectors in $V$ that are pairwise orthogonal, so that
\[ (e_j, e_l)_V = 0 \quad \text{when } j \neq l. \]
If $v \in V$, then put
\[ w = \sum_{j=1}^{n} \frac{\langle v, e_j \rangle_V}{(e_j, e_j)_V} e_j, \]
which is an element of the linear span of $e_1, \ldots, e_n$. Observe that
\[ \langle w, e_l \rangle_V = \langle v, e_l \rangle_V \]
for every $l = 1, \ldots, n$. Equivalently, $\langle w - v, e_l \rangle_V = 0$ for each $l = 1, \ldots, n$, so that
\[ \langle w - v, u \rangle_V = 0 \]
for every $u$ 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
\[ v = w, \]
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
\[ v = \sum_{j=1}^{n} \frac{\langle v, e_j \rangle_V}{(e_j, e_j)_V} e_j, \]
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
\[ \langle T(v), w \rangle_V = \langle v, T^*(w) \rangle_V \]
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
\[ T(e_l) = \sum_{j=1}^{n} \frac{\langle T(e_l), e_j \rangle_V}{(e_j, e_j)_V} e_j \]
for each $l = 1, \ldots, n$, so that
\[ \text{tr}_V T = \sum_{j=1}^{n} \frac{\langle T(e_j), e_j \rangle_V}{(e_j, e_j)_V}. \]
This implies that
\[(11.9.12) \quad \text{tr}_V T^* = \sum_{j=1}^{n} \frac{\langle T^*(e_j), e_j \rangle_V}{\langle e_j, e_j \rangle_V} = \sum_{j=1}^{n} \frac{\langle e_j, T(e_j) \rangle_V}{\langle e_j, e_j \rangle_V} = \text{tr}_V T.\]

If \(T_1, T_2 \in \mathcal{L}(V)\), then put
\[(11.9.13) \quad \langle T_1, T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (T_1 \circ T_2^*) = \text{tr}_V (T_2^* \circ T_1),\]
which defines a bilinear form on \(\mathcal{L}(V)\). Using (11.9.12), we get that
\[(11.9.14) \quad \langle T_1, T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (T_1 \circ T_2^*) = \text{tr}_V (T_2 \circ T_1^*) = \langle T_2, T_1 \rangle_{\mathcal{L}(V)},\]
as in the previous section. Observe that
\[(11.9.15) \quad \langle T, T \rangle_{\mathcal{L}(V)} = \text{tr}_V (T^* \circ T) = \sum_{j=1}^{n} \frac{\langle (T^*(T(e_j)), e_j \rangle_V}{\langle e_j, e_j \rangle_V}
= \sum_{j=1}^{n} \frac{\langle T(e_j), T(e_j) \rangle_V}{\langle e_j, e_j \rangle_V}
\]
for every \(T \in \mathcal{L}(V)\). If \(T \neq 0\), then it follows that
\[(11.9.16) \quad \langle T, T \rangle_{\mathcal{L}(V)} > 0,\]
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
\[(11.9.17) \quad \langle T_1, T_2 \rangle_{\mathcal{L}(V)} = -\text{tr}_V (T_2 \circ T_1) = -\langle T_2, T_1 \rangle_{\mathcal{L}(V)}.
\]
This means that
\[(11.9.18) \quad \langle T_1, T_2 \rangle_{\mathcal{L}(V)} = 0.\]

If \(A \in \mathcal{L}(V)\), then put
\[(11.9.19) \quad L_A(T) = A \circ T \quad \text{and} \quad R_A(T) = T \circ A\]
for every \(T \in \mathcal{L}(V)\), as before. Observe that
\[(11.9.20) \quad \langle L_A(T_1), T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (A \circ T_1 \circ T_2^*) = \text{tr}_V (T_1 \circ T_2^* \circ A) = \text{tr}_V (T_1 \circ (A^* \circ T_2)^*) = \langle T_1, L_A(T_2) \rangle_{\mathcal{L}(V)}\]
and
\[(11.9.21) \quad \langle R_A(T_1), T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (T_1 \circ A \circ T_2^*) = \text{tr}_V (T_1 \circ (T_2 \circ A^*)^*) = \langle T_1, R_A(T_2) \rangle_{\mathcal{L}(V)}\]
11.10. SELF-ADJOINT LIE SUBALGEBRAS

for every \( T_1, T_2 \in \mathcal{L}(V) \). Put

\[
(11.9.22) \quad C_A(T) = [A, T] = A \circ T - T \circ A = L_A(T) - R_A(T)
\]

for every \( T \in \mathcal{L}(V) \). Thus

\[
(11.9.23) \quad \langle C_A(T_1), T_2 \rangle_{\mathcal{L}(V)} = \langle T_1, C_A^*(T_2) \rangle_{\mathcal{L}(V)}
\]

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

\[
(11.10.1) \quad C_R(T) = [R, T] = R \circ T - T \circ R
\]

for every \( T \in gl(V) \), as in (11.9.22), which is the same as \( \text{ad}_{gl(V), R}(T) \). Remember that

\[
(11.10.2) \quad \langle C_R(T_1), T_2 \rangle_{\mathcal{L}(V)} = \langle T_1, C_R^*(T_2) \rangle_{\mathcal{L}(V)}
\]

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 \( \mathcal{A} \) be a Lie subalgebra of \( gl(V) \), as a Lie algebra over \( k \). If \( R \in \mathcal{A} \), then \( C_R \) maps \( \mathcal{A} \) into itself, and the restriction of \( C_R \) to \( \mathcal{A} \) is the same as \( \text{ad}_{\mathcal{A}, R} \). If \( R_1, R_2 \in \mathcal{A} \), then

\[
(11.10.3) \quad \text{tr}_\mathcal{A}(\text{ad}_{\mathcal{A}, R_1} \circ \text{ad}_{\mathcal{A}, R_2}) = \text{tr}_\mathcal{A}(C_{R_1} \circ C_{R_2})
\]

is the same as the Killing form on \( \mathcal{A} \) evaluated at \( R_1, R_2 \). This is the trace of \( \text{ad}_{\mathcal{A}, R_1} \circ \text{ad}_{\mathcal{A}, R_2} \) as a linear mapping from \( \mathcal{A} \) into itself, where \( \mathcal{A} \) 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 \( \mathcal{A} \), as indicated by taking the trace is taken over \( \mathcal{A} \).

If \( \mathcal{A} \) is any subset of \( \mathcal{L}(V) \), then put

\[
(11.10.4) \quad \mathcal{A}^* = \{ T^* : T \in \mathcal{A} \},
\]

which is a subset of \( \mathcal{L}(V) \) as well. Let us say that \( \mathcal{A} \) is self-adjoint as a subset of \( \mathcal{L}(V) \) when \( \mathcal{A}^* = \mathcal{A} \). Let \( \mathcal{A} \) be a Lie subalgebra of \( gl(V) \) again, and suppose that \( \mathcal{A} \) is self-adjoint as a subset of \( \mathcal{L}(V) \). If \( R \in \mathcal{A} \), then \( R^* \in \mathcal{A} \), so that \( C_R \) and \( C_{R^*} \) both map \( \mathcal{A} \) into itself. The restriction of \( C_{R^*} \) to \( \mathcal{A} \) is the same as the adjoint of the restriction of \( C_R \) to \( \mathcal{A} \), with respect to the restriction of the inner product (11.9.13) on \( \mathcal{L}(V) \) to \( \mathcal{A} \).

Let \( \mathcal{L}(\mathcal{A}) \) be the space of all linear mappings from \( \mathcal{A} \) into itself, as usual. The restriction of the inner product (11.9.13) on \( \mathcal{L}(V) \) to \( \mathcal{A} \) defines an inner
product on \( A \), as a vector space over \( k \). If \( Z \in \mathcal{L}(A) \), then let \( Z^A,* \in \mathcal{L}(A) \) be the adjoint of \( Z \) with respect to the inner product on \( A \) just mentioned. Put

\[
(Z_1, Z_2)_{\mathcal{L}(A)} = \text{tr}_A(Z_1 \circ Z_2^A,*)
\]

for every \( Z_1, Z_1 \in \mathcal{L}(A) \). This defines an inner product on \( \mathcal{L}(A) \), as a vector space over \( k \), as in the previous section.

Let \( R_1, R_2 \in A \) be given, and remember that \( \text{ad}_A R_2 \) is the same as the adjoint of \( \text{ad}_A R_2 \) with respect to the inner product on \( A \) mentioned in the preceding paragraph. It follows that

\[
\text{tr}_A(\text{ad}_A R_1 \circ \text{ad}_A R_2) = \langle \text{ad}_A R_1, \text{ad}_A R_2 \rangle_{\mathcal{L}(A)}
\]

where the right side is as in (11.10.5). In particular,

\[
\text{tr}_A(\text{ad}_A R \circ \text{ad}_A R) = \langle \text{ad}_A R, \text{ad}_A R \rangle_{\mathcal{L}(A)}
\]

for every \( R \in A \). The right side is automatically greater than or equal to 0 in \( k \), because (11.10.5) is an inner product on \( \mathcal{L}(A) \). More precisely, the right side of (11.10.7) is equal to 0 exactly when \( \text{ad}_A R = 0 \) as a linear mapping from \( A \) into itself.

### 11.11 Comparing involutions

Let \( A \) be a commutative group, where the group operations are expressed additively, and let

\[
x \mapsto x^{*,1}
\]

and

\[
x \mapsto x^{*,2}
\]

be group homomorphisms from \( A \) into itself. Suppose for the moment that these two group homomorphisms commute on \( A \), so that

\[
(x^{*,1})^{*,2} = (x^{*,2})^{*,1}
\]

for every \( x \in A \). Of course, one can define self-adjointness and anti-self-adjointness of elements of \( A \) with respect to (11.11.1) and (11.11.2) in the usual way. If \( x \in A \) is self-adjoint with respect to (11.11.2), then it follows that \( x^{*,1} \) is self-adjoint with respect to (11.11.2) as well. Similarly, if \( x \) is anti-self-adjoint with respect to (11.11.2), then \( x^{*,1} \) is anti-self-adjoint with respect to (11.11.2).

Now 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 \( ab \). Suppose that (11.11.1) is an algebra involution on \( A \), and let \( c \) be an invertible element of \( A \). Put

\[
x^{*,2} = c^{-1} x^{*,1} c
\]
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
\begin{equation}
(11.11.5) \quad c^{*,1} = c
\end{equation}
or
\begin{equation}
(11.11.6) \quad c^{*,1} = -c.
\end{equation}
In either case, (11.11.2) is an algebra involution on $\mathcal{A}$ as well, as in Section 3.14. Note that
\begin{equation}
(11.11.7) \quad c^{*,2} = c
\end{equation}
when (11.11.5) holds, and that
\begin{equation}
(11.11.8) \quad c^{*,2} = -c
\end{equation}
when (11.11.6) holds.
If $x \in \mathcal{A}$, then
\begin{equation}
(11.11.9) \quad (x^{*,1})^{*,2} = c^{-1} (x^{*,1})^{*,1} c = c^{-1} x c
\end{equation}
and
\begin{equation}
(11.11.10) \quad (x^{*,2})^{*,1} = (c^{-1} x^{*,1} c)^{*,1} = c^{*,1} (x^{*,1})^{*,1} (c^{*,1})^{-1} = c^{*,1} x (c^{*,1})^{-1}.
\end{equation}
This reduces to
\begin{equation}
(11.11.11) \quad (x^{*,2})^{*,1} = c x c^{-1}
\end{equation}
when $c$ satisfies (11.11.5) or (11.11.6). If
\begin{equation}
(11.11.12) \quad c^2 = t e
\end{equation}
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
\begin{equation}
(11.12.1) \quad \lambda_z(v) = b(v, z)
\end{equation}
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
\begin{equation}
(11.12.2) \quad \lambda_z(v) = b(v, P_0(z))
\end{equation}
for every $v \in V_0$. Equivalently, this means that
\[ b(v, z - P_0(z)) = 0 \]  
(11.12.3)
for every $v \in V$. It is easy to see that $P_0$ defines a linear mapping from $V$ into $V_0$. More precisely,
\[ P_0(z) = z \]  
(11.12.4)
when $z \in V_0$, so that $P_0$ maps $V$ onto $V_0$. It follows that
\[ P_0 \circ P_0 = P_0 \]  
(11.12.5)
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
\[ V_0 \cap Z_0 = \{0\} \]  
(11.12.6)
by (11.12.4). If $z \in V$, then
\[ z - P_0(z) \in Z_0, \]  
(11.12.7)
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
\[ b(v, z) = 0 \]  
(11.12.8)
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
\[ b(z, v) = 0 \]  
(11.12.9)
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
\[ b(P_0(u), w - P_0(w)) = b(u - P_0(u), P_0(w)) = 0, \]  
(11.12.10)
by (11.12.8) and (11.12.9). It follows that
\[ b(u, w) = b(P_0(u), P_0(w)) + b(u - P_0(u), w - P_0(w)). \]

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
\[ b(w, w) = 0 \]  
(11.12.12)
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
\[ b(x, y) \neq 0. \]  
(11.12.13)
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 on which $b(\cdot, \cdot) \equiv 0$. This additional linear subspace is not needed when $b(\cdot, \cdot)$ is nondegenerate on $V$. By construction, $b(\cdot, \cdot)$ corresponds to the bilinear form on this 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

\[(11.12.14) \quad b(x, x) \neq 0,\]

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 of $V_0$ and $Z_0$ 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,

\[(11.13.1) \quad \beta(x_j, x_l) = \beta(y_j, y_l) = 0\]

for every $j, l = 1, \ldots, n$, and

\[(11.13.2) \quad \beta(x_j, y_l) = 0\]
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 elements 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 satisfies

\[
B(x_j) = y_j, \quad B(y_j) = -x_j
\]

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
\[ \beta(T(v), w) = \beta(v, T^{*,\beta}(w)) \]
for every $v, w \in V$. This is the same as saying that
\[ \langle B(T(v)), w \rangle_V = \langle B(v), T^{*,\beta}(w) \rangle_V \]
for every $v, w \in V$, by (11.13.10). This is equivalent to
\[ \langle B(T(v)), w \rangle_V = -\langle v, B(T^{*,\beta}(w)) \rangle_V \]
for every $v, w \in V$, because of (11.13.11). This means that
\[ (B \circ T^*) = -B \circ T^{*,\beta}, \]
where the left side is the adjoint of $B \circ T$ with respect to the inner product (11.13.6). Thus
\[ T^{*,\beta} = -B^{-1} \circ (B \circ T)^* = -B^{-1} \circ T^* \circ B^* = B^{-1} \circ T^* \circ B. \]

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
\[ \beta(e_j, e_l) = 0 \]
when $j \neq l$, and for each $j = 1, \ldots, n$,
\[ \beta(e_j, e_j) \neq 0, \]
as in 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
\[ v = \sum_{j=1}^{n} v_j e_j \]
be the unique representation of $v$ as a linear combination of the $e_j$’s with coefficients $v_j \in k$. Put
\[ \langle v, w \rangle_V = \sum_{j=1}^{n} |\beta(e_j, e_j)| v_j w_j \]
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
\[ \langle v, v \rangle_V = \sum_{j=1}^{n} |\beta(e_j, e_j)| v_j^2 \]
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

\[
B(e_j) = \begin{cases} 
  e_j & \text{when } \beta(e_j, e_j) > 0 \\
  -e_j & \text{when } -\beta(e_j, e_j) > 0.
\end{cases}
\]

Clearly

\[
B^2 = I_V, \tag{11.14.7}
\]

where \( I_V \) is the identity mapping on \( V \). It is easy to see that

\[
B^* = B, \tag{11.14.8}
\]

where \( B^* \) is the adjoint of \( B \) with respect to the inner product (11.14.4). Of course,

\[
\langle B(v), w \rangle_V = \sum_{j=1}^{n} \beta(e_j, e_j) v_j w_j \tag{11.14.9}
\]

for every \( v, w \in V \), by construction. One can check that

\[
\beta(v, w) = \langle B(v), w \rangle_V \tag{11.14.10}
\]

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

\[
\beta(T(v), w) = \beta(v, T^{*,\beta}(w)) \tag{11.14.11}
\]

for every \( v, w \in V \), as before. This means that

\[
\langle B(T(v)), w \rangle_V = \langle B(v), T^{*,\beta}(w) \rangle_V \tag{11.14.12}
\]

for every \( v, w \in V \), because of (11.14.10). This is the same as saying that

\[
\langle B(T(v)), w \rangle_V = \langle v, B(T^{*,\beta}(w)) \rangle_V \tag{11.14.13}
\]

for every \( v, w \in V \), by (11.14.8). This is equivalent to asking that

\[
(B \circ T)^* = B \circ T^{*,\beta} \tag{11.14.14}
\]

where the left side is the adjoint of \( B \circ T \) with respect to the inner product (11.14.4). This shows that

\[
T^{*,\beta} = B^{-1} \circ (B \circ T)^* = B^{-1} \circ T^* \circ B^* = B^{-1} \circ T^* \circ B. \tag{11.14.15}
\]
Chapter 12

Some complex versions

12.1 Complexifying ordered fields

Let \( k \) be a 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

\[
z = x + y \sqrt{-a},
\]

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

\[
w = u + v \sqrt{-a}
\]

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

\[
z + w = (x + u) + (y + v) \sqrt{-a}
\]

and

\[
z w = (x u - y v a) + (x v + y u) \sqrt{-a},
\]

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

\[
\overline{z} = x - y \sqrt{-a}.
\]

One can verify that

\[
\overline{z + w} = \overline{z} + \overline{w}
\]

as usual.
and
\[(12.1.7) \quad \overline{z \cdot w} = \overline{z} \cdot \overline{w}\]
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
\[(12.1.9) \quad z + \overline{z} = 2 : x\]
and
\[(12.1.10) \quad z \cdot \overline{z} = x^2 + a y^2\]
correspond to elements of \(k\). Note that
\[(12.1.11) \quad z \cdot \overline{z} > 0\]
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 \(z \cdot \overline{z}\) 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 \(\frac{1}{z \cdot \overline{z}}\). 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
\[(12.1.12) \quad T(\sqrt{-a} v) = \sqrt{-a} \cdot T(v)\]
for every \(v \in V\). A \(k\)-linear mapping \(T\) from \(V\) into \(W\) is said to be conjugate-linear if
\[(12.1.13) \quad T(\sqrt{-a} v) = -\sqrt{-a} \cdot T(v)\]
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 \textit{sesquilinear} if \(b(v, w)\) is \(k[\sqrt{-a}]\)-linear in \(v\) for every \(w \in V\), 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 \textit{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 \textit{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 \textit{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 \textit{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.
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

$$b(T(v), w) = b(v, T^*(w)) \quad (12.2.6)$$

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

$$T_2 \circ T_1^* = T_1^* \circ T_2^* \quad (12.2.7)$$

as usual. If $b(\cdot, \cdot)$ is Hermitian-symmetric on $V$, then one can verify that

$$(T^*)^* = T \quad (12.2.8)$$

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

$$\langle v, v \rangle_V \in k \quad (12.3.1)$$

for every $v \in V$, because $\overline{\langle v, v \rangle_V} = \langle v, v \rangle_V$, by Hermitian symmetry. Suppose that

$$\langle v, v \rangle_V > 0 \quad (12.3.2)$$

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

$$\langle v, w \rangle_V = 0, \quad (12.3.3)$$

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

$$w = \sum_{j=1}^n \frac{\langle v, e_j \rangle_V}{\langle e_j, e_j \rangle_V} e_j \quad (12.3.4)$$

is an element of the linear span of $e_1, \ldots, e_n$ in $V$, and

$$\langle w, e_i \rangle_V = \langle v, e_i \rangle_V \quad (12.3.5)$$
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
\[
(12.3.6) \quad \langle w - v, u \rangle_V = 0
\]
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
\[
(12.3.7) \quad v = \sum_{j=1}^{n} \frac{\langle v, e_j \rangle_V}{\langle e_j, e_j \rangle_V} e_j,
\]
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
\[
(12.3.8) \quad \langle T(v), w \rangle_V = \langle v, T^*(w) \rangle_V
\]
for every \( v, w \in V \), as before. Of course,
\[
(12.3.9) \quad T(e_l) = \sum_{j=1}^{n} \frac{\langle T(e_l), e_j \rangle_V}{\langle e_j, e_j \rangle_V} e_j
\]
for every \( l = 1, \ldots, n \), as in (12.3.7). This implies that
\[
(12.3.10) \quad \text{tr}_V T = \sum_{j=1}^{n} \frac{\langle T(e_j), e_j \rangle_V}{\langle e_j, e_j \rangle_V}.
\]
Thus
\[
(12.3.11) \quad \text{tr}_V T^* = \sum_{j=1}^{n} \frac{\langle T^*(e_j), e_j \rangle_V}{\langle e_j, e_j \rangle_V} = \sum_{j=1}^{n} \frac{\langle e_j, T(e_j) \rangle_V}{\langle e_j, e_j \rangle_V} = \sum_{j=1}^{n} \frac{\langle T(e_j), e_j \rangle_V}{\langle e_j, e_j \rangle_V} = \text{tr}_V T.
\]

Put
\[
(12.3.12) \quad \langle T_1, T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (T_1 \circ T_2^*) = \text{tr}_V (T_2^* \circ T_1)
\]
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
\[
(12.3.13) \quad \langle T_1, T_2 \rangle_{\mathcal{L}(V)} = \text{tr}_V (T_1 \circ T_2^*) = \text{tr}_V (T_2 \circ T_1)^* = \langle T_2, T_1 \rangle_{\mathcal{L}(V)}
\]
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^*(T(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}.
\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.

\section*{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 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 $ad_{\mathcal{A}, R}$. Let $R_1, R_2 \in \mathcal{A}$ be given, so that

$$\text{tr}_{\mathcal{A}}(ad_{\mathcal{A}, R_1} \circ ad_{\mathcal{A}, R_2}) = \text{tr}_{\mathcal{A}}(C_{R_1} \circ C_{R_2})$$

(12.4.3)

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

$$\mathcal{A}^* = \{ T^* : T \in \mathcal{A} \},$$

(12.4.4)

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 $Z^{\mathcal{A}^*} \in L(\mathcal{A})$ be the adjoint of $Z$ with respect to this inner product. If $Z_1, Z_2 \in L(\mathcal{A})$, then put

$$\langle Z_1, Z_2 \rangle_{L(\mathcal{A})} = \text{tr}_{\mathcal{A}}(Z_1 \circ Z_2^{\mathcal{A}^*}).$$

(12.4.5)

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 $ad_{\mathcal{A}, R_2^*}$ is the same as the adjoint of $ad_{\mathcal{A}, R_2}$ with respect to the inner product on $\mathcal{A}$ mentioned in the preceding paragraph. This implies that

$$\text{tr}_{\mathcal{A}}(ad_{\mathcal{A}, R_1} \circ ad_{\mathcal{A}, R_2}) = \langle ad_{\mathcal{A}, R_1}, ad_{\mathcal{A}, R_2} \rangle_{L(\mathcal{A})},$$

(12.4.6)

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

$$\text{tr}_{\mathcal{A}}(ad_{\mathcal{A}, R} \circ ad_{\mathcal{A}, R^*}) = \langle ad_{\mathcal{A}, R}, ad_{\mathcal{A}, R} \rangle_{L(\mathcal{A})}.$$

(12.4.7)

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 $ad_{\mathcal{A}, R} \neq 0$ 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
\begin{equation}
(12.5.11) \quad b(x, x) \neq 0,
\end{equation}
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
\begin{equation}
(12.5.12) \quad b(w, w) = 0
\end{equation}
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
\begin{equation}
(12.6.1) \quad \beta(e_j, e_l) = 0
\end{equation}
when $j \neq l$, and for each $j = 1, \ldots, n$,
\begin{equation}
(12.6.2) \quad \beta(e_j, e_j) \neq 0,
\end{equation}
as in the previous section. Note that
\begin{equation}
(12.6.3) \quad \beta(e_j, e_j) \in k
\end{equation}
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
\begin{equation}
(12.6.4) \quad v = \sum_{j=1}^{n} v_j e_j,
\end{equation}
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

\[
\langle v, w \rangle_V = \sum_{j=1}^{n} |\beta(e_j, e_j)| v_j \overline{w_j},
\]

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

\[
\beta(v, w) = \langle B(v), w \rangle_V
\]

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 \( T^{*,\beta} \in \mathcal{L}(V) \) of \( T \) with respect to \( \beta(\cdot, \cdot) \) is characterized by the condition that

\[
\beta(T(v), w) = \beta(v, T^{*,\beta}(w))
\]

(12.6.9)

for every \( v, w \in V \). This is the same as saying that

\[
\langle B(T(v)), w \rangle_V = \langle B(v), T^{*,\beta}(w) \rangle_V
\]

(12.6.10)

for every \( v, w \in V \), by (12.6.8). Equivalently, this means that

\[
\langle B(T(v)), w \rangle_V = \langle v, B(T^{*,\beta}(w)) \rangle_V
\]

(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

\[
(B \circ T)^* = B \circ T^{*,\beta},
\]

(12.6.12)

where the left side is the adjoint of \( B \circ T \) with respect to the inner product (12.6.5). Thus

\[
T^{*,\beta} = B^{-1} \circ (B \circ T)^* = B^{-1} \circ T^* \circ B^* = B^{-1} \circ T^* \circ B.
\]

(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

\[(12.7.1)\]

$$\text{Re}_{k[\sqrt{-a}]} z = x = (1/2) (z + \overline{z})$$

the real part of $z$. Note that

\[(12.7.2)\]

$$\text{Re}_{k[\sqrt{-a}]}(\sqrt{-a} z) = -a y = (\sqrt{-a}/2) (z - \overline{z}).$$

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

\[(12.7.3)\]

$$\langle v, w \rangle_{V_k} = \text{Re}_{k[\sqrt{-a}]}(\langle v, w \rangle_{V_k[\sqrt{-a}]})$$

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

\[(12.7.4)\]

$$\langle T(v), w \rangle_{V_k[\sqrt{-a}]} = \langle v, T^*_{V_k[\sqrt{-a}]}(w) \rangle_{V_k[\sqrt{-a}]}$$

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

\[(12.7.5)\]

$$\langle T(v), w \rangle_{V_k} = \langle v, T^*_{V_k}(w) \rangle_{V_k}$$

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

\[(12.7.6)\]

$$T^*_{V_k} = T^*_{V_k[\sqrt{-a}]}.$$

Put

\[(12.7.7)\]

$$J_a(v) = \sqrt{-a} v$$

for every $v \in V$, which defines a $k[\sqrt{-a}]$-linear mapping from $V$ into itself. By construction,

\[(12.7.8)\]

$$J^2_a = -a I,$$

where $I$ is the identity mapping on $V$, and

\[(12.7.9)\]

$$\langle J_a(v), w \rangle_{V_k} = \text{Re}_{k[\sqrt{-a}]}(\sqrt{-a} \langle v, w \rangle_{V_k[\sqrt{-a}]})$$

for every $v, w \in V$. One can check that

\[(12.7.10)\]

$$J^*_{V_k} = J^*_{V_k[\sqrt{-a}]} = -J_a.$$
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

\[ \beta_{V_k}(J_a(v), w) = -\beta_{V_k}(v, J_a(w)) = \text{Re}_{k[\sqrt{-a}]}(\sqrt{-a} \beta_{V_C}(v, w)) \]

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 $V$, and (12.7.12) is antisymmetric on $V$. Of course, (12.7.11) and (12.7.12) correspond to (12.7.3) and (12.7.9) when $\beta_{V_C}(\cdot, \cdot) = \langle \cdot, \cdot \rangle_{V_C}$.

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 $T^* \beta_{V_{k[\sqrt{-a}]}} \in \mathcal{L}_{k[\sqrt{-a}]}(V_{k[\sqrt{-a}]})$ such that

\[ \beta_{V_{k[\sqrt{-a}]}}(T(v), w) = \beta_{V_{k[\sqrt{-a}]}}(v, T^* \beta_{V_{k[\sqrt{-a}]}}(w)) \]

for every $v, w \in V$, as usual. Similarly, if $T \in \mathcal{L}_k(V_k)$, then there is a unique $T^* \beta_{V_k} \in \mathcal{L}_k(V_k)$ such that

\[ \beta_{V_k}(T(v), w) = \beta_{V_k}(v, T^* \beta_{V_k}(w)) \]

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

\[ T^* \beta_{V_k} = T^* \beta_{V_C}, \]

as before. In particular, $J_a^* \beta_{V_k} = J_a^* \beta_{V_C} = -J_a$.

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 $\beta_{V_C}(\cdot, \cdot)$, 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

$\gamma_{V_k}(v, w) = \text{Re}_{k[\sqrt{-a}]}(\gamma_{V_{k[\sqrt{-a}]}}(v, w))$

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

$\gamma_{V_k}(J_a(v), w) = \gamma_{V_k}(v, J_a(w)) = \text{Re}_{k[\sqrt{-a}]}(\sqrt{-a} \gamma_{V_{k[\sqrt{-a}]}}(v, w))$

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

$\alpha_{V_k}(J_a(v), w) = \gamma_{V_k}(v, J_a(w))$

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 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

$\gamma_{V_{k[\sqrt{-a}]}}(T(v), w) = \gamma_{V_{k[\sqrt{-a}]}}(v, T^{*, \gamma_{V_{k[\sqrt{-a]}}}}(w))$

for every $v, w \in V$. Similarly, if $T \in \mathcal{L}_k(V_k)$, then there is a unique $T^{*, \gamma_{V_k}}$ in $\mathcal{L}_k(V_k)$ such that

$\gamma_{V_k}(T(v), w) = \gamma_{V_k}(v, T^{*, \gamma_{V_k}}(w))$

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

$T^{*, \gamma_{V_k}} = T^{*, \gamma_{V_{k[\sqrt{-a}]}}}$.

Note that

$J_a^{*, \gamma_{V_k}} = J_a^{*, \gamma_{V_{k[\sqrt{-a}]}}} = J_a$. 

Let \( \langle \cdot, \cdot \rangle_{V_k(\sqrt{-a})} \) 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

\[
\gamma_{V_k(\sqrt{-a})}(v, w) = \langle v, C(w) \rangle_{V_k(\sqrt{-a})}
\]

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 \( \langle \cdot, \cdot \rangle_{V_k(\sqrt{-a})} \) is conjugate-linear in the second variable.

Suppose that \( \gamma_{V_k(\sqrt{-a})}(\cdot, \cdot) \) 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

\[
\langle T(v), C(w) \rangle_{V_k(\sqrt{-a})} = \langle v, C(T^*\gamma_{V_k(\sqrt{-a})}(w)) \rangle_{V_k(\sqrt{-a})}
\]

for every \( v, w \in V \), by (12.8.5) and (12.8.9). We also have that

\[
\langle T(v), C(w) \rangle_{V_k(\sqrt{-a})} = \langle v, T^*\gamma_{V_k(\sqrt{-a})}(C(w)) \rangle_{V_k(\sqrt{-a})}
\]

for every \( v, w \in V \), where \( T^*\gamma_{V_k(\sqrt{-a})} \) is the adjoint of \( T \) with respect to the inner product \( \langle \cdot, \cdot \rangle_{V_k(\sqrt{-a})} \) on \( V_{k[\sqrt{-a}]} \). It follows that

\[
C \circ T^*\gamma_{V_k(\sqrt{-a})} = T^*\gamma_{V_k(\sqrt{-a})} \circ C.
\]

### 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 \( \langle \cdot, \cdot \rangle_{V_k(\sqrt{-a})} \) 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 \( C^*\gamma_{V_k} \) 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

\[
(12.9.1) \quad C \circ J_a = -J_a \circ C,
\]
12.10. ANTISYMMETRIC FORMS OVER $K[\sqrt{-A}]$

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

\[(12.9.5)\quad \gamma_{V_k[\sqrt{-a}]}(v, w) = \langle v, C(w) \rangle_{V_k[\sqrt{-a}]},\]

for every $v, w \in V$. This defines a bilinear form on $V_{k[\sqrt{-a}]}$, because $C$ is conjugate-linear, and $\langle \cdot, \cdot \rangle_{V_k[\sqrt{-a}]}$ 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

\[(12.9.7)\quad \gamma_{V_k[\sqrt{-a}]}(v, w) = \langle w, C^{*,V_k}(v) \rangle_{V_k[\sqrt{-a}]},\]

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
\[
(v, w)_V = \sum_{j=1}^{n} v_{x_j} w_{x_j} + \sum_{j=1}^{n} v_{y_j} w_{y_j}
\]
for every \( v, w \in V \). 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
\[
C(v) = -\sum_{j=1}^{n} \overline{v_j} y_j + \sum_{j=1}^{n} \overline{w_j} x_j.
\]
One can verify that
\[
\gamma(v, w) = (v, C(w))_V
\]
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 \( (v, w)_{V_k} \) 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 \( (\cdot, \cdot)_{V_k} \). 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

(12.11.1) $\gamma(e_j, e_l) = 0$

when $j \neq l$, and for each $j = 1, \ldots, n$,

(12.11.2) $\gamma(e_j, e_j) \neq 0$

and

(12.11.3) $\gamma(e_j, e_j) \in k$.

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

(12.11.4) $\gamma(e_j, e_j) > 0$

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

(12.11.5) $\gamma(e_j, e_j) = 1$

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

(12.11.6) $v = \sum_{j=1}^{n} v_j e_j,$

where $v_j \in k[\sqrt{-a}]$ for each $j = 1, \ldots, n$. If $v, w \in V$, then put

(12.11.7) $\langle v, w \rangle_V = \sum_{j=1}^{n} |\gamma(e_j, e_j)| v_j w_j,$

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

(12.11.8) $C(e_j) = e_j$ when $\gamma(e_j, e_j) > 0$

$= -e_j$ when $-\gamma(e_j, e_j) > 0$. 

If (12.11.4) holds, then $C(e_j) = e_j$ for every $j = 1, \ldots, n$, so that
\begin{equation}
C(v) = \sum_{j=1}^{n} \gamma(e_j) v_j e_j 
\end{equation}
for every $v \in V$ as in (12.11.6). Observe that
\begin{equation}
\langle v, C(w) \rangle_V = \sum_{j=1}^{n} \gamma(e_j) v_j w_j 
\end{equation}
for every $v, w \in V$. Using this, one can check that
\begin{equation}
\gamma(v, w) = \langle v, C(w) \rangle_V 
\end{equation}
for every $v, w \in V$. One can also verify that
\begin{equation}
C^2 = I_V, 
\end{equation}
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
\begin{equation}
\langle (T^* \circ T)(v), v \rangle_V = \langle T(v), T(v) \rangle_W = \| T(v) \|_W^2 
\end{equation}
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 $\langle \cdot, \cdot \rangle$, so that
\begin{equation}
\langle A(v), w \rangle = \langle v, A(w) \rangle 
\end{equation}
for every $v, w \in V$. In the complex case, (12.12.2) implies that
\begin{equation}
\langle A(v), v \rangle = \langle v, A(v) \rangle = \langle A(v), v \rangle 
\end{equation}
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

\[(12.12.4) \quad \langle A(v), v \rangle \geq 0\]

for every $v \in V$. Similarly, $A$ is said to be strictly positive on $V$ if

\[(12.12.5) \quad \langle A(v), v \rangle_V > 0\]

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

\[(12.13.1) \quad R^2 = T^* \circ T,\]

as in the previous section. If $u, v \in V$, then

\[(12.13.2) \quad \langle (T^* \circ T)(u), v \rangle_V = \langle T(u), T(v) \rangle_W\]
CHAPTER 12. SOME COMPLEX VERSIONS

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 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

\[
\beta(v, w) = \langle B(v), w \rangle \tag{12.14.1}
\]
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

\[
A^2 = B \circ B^*. \tag{12.14.2}
\]

More precisely, \(A\) is invertible on \(V\), so that \(A\) is strictly positive on \(V\). Put

\[
B_0 = A^{-1} \circ B, \tag{12.14.3}
\]

which is invertible on \(V\) as well. By construction,

\[
B_0 \circ B_0^* = A^{-1} \circ B \circ B^* \circ A^{-1} = I, \tag{12.14.4}
\]

so that

\[
B_0^{-1} = B_0^*. \tag{12.14.5}
\]

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

\[
B_0^2 = A^{-2} \circ B^2 = I. \tag{12.14.6}
\]

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

\[
B_0^2 = A^{-2} \circ B^2 = -I. \tag{12.14.7}
\]

Alternatively, one can verify that \(B_0\) is anti-self-adjoint in this case.

Clearly

\[
B = A \circ B_0, \tag{12.14.8}
\]

by (12.14.3). Put

\[
\langle v, w \rangle_A = \langle A(v), w \rangle \tag{12.14.9}
\]
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

$$\beta(v, w) = \langle B_0(v), w \rangle_A$$

(12.14.10)

for every $v, w \in V$. If $B$ is normal with respect to $\langle \cdot, \cdot \rangle$ on $V$, then

$$\langle B_0(v), w \rangle_A = \langle A(B_0(v)), w \rangle = \langle A(v), B_0^*(w) \rangle = \langle v, B_0^*(w) \rangle_A$$

(12.14.11)

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 $B$ is a 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 $V$. 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

$$B_0^* = I$$

(12.14.12)

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

$$\gamma(v, w) = \langle v, C(w) \rangle$$

(12.15.1)

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 $V_{\mathbb{R}}$ be $V$ considered as a vector space over the real numbers, and let $\langle \cdot, \cdot \rangle_{V_{\mathbb{R}}}$ be the real part of $\langle \cdot, \cdot \rangle$, which is an inner product on $V_{\mathbb{R}}$. Also let $C$ be a conjugate-linear mapping from $V$ into itself, and let $C^*:V_{\mathbb{R}}$ be the adjoint
of \(C\), as a real-linear mapping from \(V\) into itself, with respect to \(\langle \cdot, \cdot \rangle_{VR}\). Thus \(C^{*,VR}\) is conjugate-linear on \(V\) too, as in Section 12.9. It follows that \(C \circ C^{*,VR}\) is a complex-linear mapping from \(V\) into itself, as before.

Note that \(C \circ C^{*,VR}\) is nonnegative and self-adjoint with respect to \(\langle \cdot, \cdot \rangle_{VR}\). This implies that there is a unique real-linear mapping \(A\) from \(V\) into itself that is nonnegative and self-adjoint with respect to \(\langle \cdot, \cdot \rangle_{VR}\) and satisfies

\[
(12.15.2) \quad A^2 = C \circ C^{*,VR}.
\]

Remember that \(A\) commutes with any real-linear mapping from \(V\) into itself that commutes with \(C \circ C^{*,VR}\). In particular, \(A\) commutes with the mapping \(J\) from \(V\) into itself that corresponds to multiplication by \(i\), because \(C \circ C^{*,VR}\) 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 \(\langle \cdot, \cdot \rangle_{VR}\).

Alternatively, \(C \circ C^{*,VR}\) is self-adjoint as a complex-linear mapping from \(V\) into itself with respect to \(\langle \cdot, \cdot \rangle_{VR}\), because \(C \circ C^{*,VR}\) is nonnegative as a self-adjoint real-linear mapping from \(V\) into itself with respect to \(\langle \cdot, \cdot \rangle_{VR}\). Similarly, it is easy to see that \(C \circ C^{*,VR}\) is nonnegative as a self-adjoint real-linear mapping from \(V\) into itself with respect to \(\langle \cdot, \cdot \rangle_{VR}\). 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 \(\langle \cdot, \cdot \rangle_{VR}\) 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 \(C^{*,VR}\) is invertible on \(V\) too, and that \(C \circ C^{*,VR}\) is strictly positive on \(V\) with respect to \(\langle \cdot, \cdot \rangle_{VR}\), and hence with respect to \(\langle \cdot, \cdot \rangle\). It follows that \(A\) is invertible on \(V\), and strictly positive with respect to \(\langle \cdot, \cdot \rangle_{VR}\) and \(\langle \cdot, \cdot \rangle\). Thus

\[
(12.15.3) \quad \langle v, w \rangle_{VR,A} = \langle A(v), w \rangle_{VR}
\]

defines an inner product on \(VR\), and

\[
(12.15.4) \quad \langle v, w \rangle_A = \langle A(v), w \rangle
\]

defines an inner product on \(V\). Of course,

\[
(12.15.5) \quad \langle v, w \rangle_{VR,A} = \text{Re} \langle v, w \rangle_A
\]

for every \(v, w \in V\), by construction.

Put

\[
(12.15.6) \quad C_0 = A^{-1} \circ C,
\]

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, \(C_0 \circ C^{*,VR} = I\), so that

\[
(12.15.7) \quad C_0^{-1} = C^{*,VR}.
\]
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 $C^{\ast, V_R}$. This implies that $C$ commutes with $C \circ C^{\ast, V_R}$, and hence that $A$ commutes with $C$. It follows that $C_0$ commutes with $A$ and $C$ in this situation. Thus

$$
\langle C_0(v), w \rangle_{V_{R,A}} = \langle A(C_0(v)), w \rangle_{V_R} = \langle C_0(A(v)), w \rangle_{V_R}
$$

(12.15.9)

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 $C^{\ast, V_R}_0$ in this case.

If $\gamma(\cdot, \cdot)$ is symmetric or antisymmetric on $V$, then

$$
\gamma_R(v, w) = \text{Re} \gamma(v, w) = \langle v, C(w) \rangle_{V_R}
$$

(12.15.10)

has the analogous property as a bilinear form on $V_R$. This means that $C$ is self-adjoint or antiself-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

$$
C^2_0 = I,
$$

(12.15.11)

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

$$
C^2_0 = -I.
$$

(12.15.12)
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 $(A, [·, ·]_A)$ 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

$$\beta(u_j, w_l) = \delta_{j,l}$$

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
\begin{equation}
(13.2.7)
a_{j,r} = -b_{r,j}
\end{equation}
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
\begin{equation}
(13.2.8)
c_{\rho}(\beta) = \sum_{j=1}^{n} \rho_{u_{j}} \circ \rho_{w_{j}},
\end{equation}
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
\begin{equation}
(13.2.9)
[\rho_{x}, c_{\rho}(\beta)] = \sum_{j=1}^{n} [\rho_{x}, \rho_{u_{j}} \circ \rho_{w_{j}}],
\end{equation}
using the commutator bracket on $\mathcal{L}(V)$ corresponding to composition of linear mappings on $V$. It follows that
\begin{equation}
(13.2.10)
[\rho_{x}, c_{\rho}(\beta)] = \sum_{j=1}^{n} \rho_{u_{j}} \circ [\rho_{x}, \rho_{w_{j}}],
\end{equation}
as in Section 2.5. Because $\rho$ is a Lie algebra representation, we get that
\begin{equation}
(13.2.11)
[\rho_{x}, c_{\rho}(\beta)] = \sum_{j=1}^{n} \rho_{u_{j}} \circ [\rho_{x}, \rho_{w_{j}}],
\end{equation}
using (13.2.3) and (13.2.4) in the second step, and (13.2.7) in the third step.

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
\begin{equation}
(13.2.12)
u'_{h} = \sum_{j=1}^{n} \mu_{h,j} u_{j}
\end{equation}
for every $h = 1, \ldots, n$ and
\begin{equation}
(13.2.13)
u'_{r} = \sum_{l=1}^{n} \nu_{r,l} w_{l}
\end{equation}
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
\begin{equation}
(13.2.14)
\beta(u'_{h}, w'_{r}) = \sum_{j=1}^{n} \sum_{l=1}^{n} \mu_{h,j} \nu_{r,l} \beta(u_{j}, w_{l})
\end{equation}
\begin{equation}
= \sum_{j=1}^{n} \sum_{l=1}^{n} \mu_{h,j} \nu_{r,l} \delta_{j,l} = \sum_{j=1}^{n} \mu_{h,j} \nu_{r,j}
\end{equation}
for every $h, r = 1, \ldots, n$, using (13.2.2) in the second step. It follows that

\begin{equation}
(13.2.15) \hspace{1cm} \beta(u'_h, w'_r) = \delta_{h,r}
\end{equation}

for every $h, r = 1, \ldots, n$ if and only if

\begin{equation}
(13.2.16) \hspace{1cm} \sum_{j=1}^{n} \mu_{h,j} \nu_{r,j} = \delta_{h,r}
\end{equation}

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

\begin{equation}
(13.2.17) \hspace{1cm} \sum_{h=1}^{n} \nu_{h,l} \mu_{h,j} = \delta_{l,j}
\end{equation}

for every $j, l = 1, \ldots, n$. If $\rho$ is as in the previous paragraph, then

\begin{equation}
(13.2.18) \hspace{1cm} \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}.
\end{equation}

If (13.2.17) holds, then we get that

\begin{equation}
(13.2.19) \hspace{1cm} \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},
\end{equation}

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

\begin{equation}
(13.3.1) \hspace{1cm} \beta_z(x) = \beta(x, z)
\end{equation}
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 respresentation of \( A \) on \( \mathcal{L}(A) \). 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 \textit{Casimir element} associated to \( \beta \), 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 \( \mathcal{L}(V) \), and thus a linear mapping from \( A \otimes A \) into \( \mathcal{L}(V) \). 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

\[
\sum_{h=1}^{n} u'_h \otimes w'_h = \sum_{h=1}^{n} \sum_{j=1}^{n} \sum_{l=1}^{n} \mu_{h,j} v_{h,l} u_j \otimes w_l.
\]

If (13.2.17) holds, then it follows that

\[
\sum_{h=1}^{n} u'_h \otimes w'_h = \sum_{j=1}^{n} \sum_{l=1}^{n} \delta_{j,l} u_j \otimes w_l = \sum_{j=1}^{n} u_j \otimes w_j.
\]

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 7.14. 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 $\text{gl}(V)$.

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 $\text{gl}(V)$. Using Cartan’s criterion, we get that (13.4.3) is solvable as a Lie algebra over $k$. This implies that (13.4.2) is solvable, 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 $\mathcal{L}(V)$ 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
13.5. A SPLITTING THEOREM

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

\[
\beta_{\rho^V}(x, y) = \text{tr}_V(\rho^V_x \circ \rho^V_y)
\]

and

\[
\beta_{\rho^W}(x, y) = \text{tr}_W(\rho^W_x \circ \rho^W_y),
\]

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 \( gl(V) \). 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 \( gl(W) \). 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 \( \rho^V_y = 0 \) 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 \( gl(W) \), since otherwise we could replace \( A \) with \( A/A_0 \).
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 $c_{\rho^W} = c_{\rho^W}(\beta_{\rho^W})$ 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

\begin{equation}
\beta_z(x) = \beta(x, z)
\end{equation}

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

\begin{equation}
T_{y,z}(x) = \beta_z(x)y = \beta(x, z)y
\end{equation}

for every $x \in A$, which defines $T_{y,z}$ as a linear mapping from $A$ into itself. Observe that

\begin{equation}
\beta(T_{y,z}(x), w) = \beta(x, z)\beta(y, w)
\end{equation}

and

\begin{equation}
\beta(x, T_{z,y}(w)) = \beta(x, z)\beta(w, y)
\end{equation}

for every $w, x, y, z \in A$. If $\beta(\cdot, \cdot)$ is symmetric as a bilinear form on $A$, then we get that

\begin{equation}
\beta(T_{y,z}(x), w) = \beta(x, T_{z,y}(w))
\end{equation}

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

\begin{equation}
\beta(T(x), w) = \beta(x, T^*(w))
\end{equation}

for every $w, x \in A$, as in Section 2.14. If $\beta(\cdot, \cdot)$ is symmetric on $A$, then

\begin{equation}
(T_{y,z})^* = T_{z,y}
\end{equation}
CHAPTER 13. SEMISIMPLICITY

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 7.14. 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 7.14. 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 \( \text{gl}(A) \), 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

\[
(13.8.1) \quad Z(A) \cap B = \{0\}, \quad Z(A) + B = A,
\]

and that \( B \) is an ideal in \( A \), as before.

Observe that

\[
(13.8.2) \quad [A, A] = [B, B] \subseteq B.
\]

We also have that

\[
(13.8.3) \quad [A/Z(A), A/Z(A)] = A/Z(A),
\]

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

\[
(13.8.4) \quad B = [A, A],
\]

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

$$[a, [w, y]_B]_B = 0,$$

(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

\[(13.10.2) \quad b(x, y) = \text{tr}_B(\text{ad}_{B,x} \circ \text{ad}_{B,y})\]

for every $x, y \in B$, which is the Killing form on $B$, as in Section 7.9. Remember that

\[(13.10.3) \quad b([x, w]_B, y) = b(x, [w, y]_B)\]

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

\[(13.10.4) \quad b_A(x, y) = \text{tr}_A(\text{ad}_{A,x} \circ \text{ad}_{A,y})\]

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

\[(13.10.5) \quad b_A(x, y) = b(x, y),\]

as in Section 7.10, because $A$ is an ideal in $B$.

Put

\[(13.10.6) \quad A^\perp = \{ x \in B : b(x, y) = 0 \text{ for every } y \in A \}.\]

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

\[(13.10.7) \quad A \cap A^\perp = \{0\}\]

in this situation. Note that $[A, A^\perp] \subseteq A \cap A^\perp$, because $A$ and $A^\perp$ are ideals in $A$. Hence

\[(13.10.8) \quad [A, A^\perp] = \{0\},\]

by \((13.10.7)\).

If $x \in B$ and $y \in A$, then put

\[(13.10.9) \quad \lambda_x(y) = b(x, y),\]

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

\[(13.10.10) \quad \lambda_x(y) = b_A(w, y)\]

for every $y \in A$. This means that

\[(13.10.11) \quad b(x, y) = b_A(w, y) = b(w, y)\]

for every $y \in A$, so that $x - w \in A^\perp$. Thus

\[(13.10.12) \quad A + A^\perp = B.\]
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

\[
\text{ad} A = \{ \text{ad} x : x \in A \}
\]

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

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

(13.11.2)

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

This means that

\[
[\delta, \text{ad} x] = \text{ad} \delta(x),
\]

(13.11.3)

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

Nilpotency and diagonalizability

14.1 A duality argument

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 put $ad_x(y) = [x, y]_A$ 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

\[(14.1.1) \quad \beta(ad_x(w), z) = -\beta(w, ad_x(z))\]

for every $w, x, z \in A$. In particular,

\[(14.1.2) \quad \beta(ad_x(w), z) = 0\]

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

\[(14.1.3) \quad \{ad_x(w) : w \in A\}\]

of $ad_x$ is a linear subspace of $A$. If $z \in A$, then (14.1.2) holds for every $w \in A$ if and only if

\[(14.1.4) \quad \beta(w, ad_x(z)) = 0\]

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 $ad_x(z) = 0$. Thus (14.1.2) holds for every $w \in A$ if and only if $ad_x(z) = 0$.

Consider

\[(14.1.5) \quad \{y \in A : \beta(y, z) = 0 \text{ for every } z \in A \text{ such that } [x, z]_A = 0\},\]
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 $\text{ad}_z$ 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 nilpotency

Let $k$ be a field of characteristic 0, and let $\mathcal{A}$ be an associative algebra over $k$, where multiplication of $a, b \in \mathcal{A}$ is expressed as $ab$. Also let $\delta$ be a derivation on $\mathcal{A}$, and suppose that $b \in \mathcal{A}$ 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 $\mathcal{A}$, as a vector space over $k$. If $\mathcal{A}$ has finite dimension as a vector space over $k$, then $b^j = 0$ for some $j \geq 1$.

Remember that

\[
\delta_a(x) = [a, x] = ax - xa
\]
defines a derivation on \( A \) for every \( a \in A \). Suppose that \( b \in A \) is an eigenvector of \( \delta_a \) with eigenvalue \( \lambda \in k \), which is to say that
\[
\delta_a(b) = \mathbb{a, b} = \lambda b.
\]
If \( \lambda \neq 0 \) and \( A \) has finite dimension as a vector space over \( k \), then \( b \) is nilpotent in \( A \), as in the preceding paragraph.

Now let \(( A, [\cdot, \cdot]_A) \) be a Lie algebra over \( k \), and suppose that \( u, x \in A \) satisfy
\[
[u, x]_A = x.
\]
(14.2.5)

Let \( V \) be a vector space over \( k \), and let \( \rho \) be a representation of \( A \) as a Lie algebra on \( V \). Thus
\[
\rho[u, x]_A = \rho u \circ \rho x - \rho x \circ \rho u
\]
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 \( k \) 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 \( A \) has finite dimension as a vector space over \( k \). Let \( x \) be an ad-nilpotent element of \( A \), so that \( \text{ad}_x \) is nilpotent on \( A \). If \( A \) is semisimple as a Lie algebra over \( k \), then the Killing form on \( A \) is nondegenerate, as in Section 10.13. In this case, there is a \( u \in A \) 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 \( k \) be a field, let \( V \) be a finite-dimensional vector space over \( k \), and let \( T \) be a linear mapping from \( V \) into itself. If \( k \) is algebraically closed and \( T \) is diagonalizable on \( V \), then one may say that \( T \) is semisimple as a linear mapping on \( V \), as on p17 of [13] and p40 of [24]. Otherwise, one may say that \( T \) is semisimple on \( V \) when \( T \) becomes diagonalizable after passing to an algebraic closure of \( k \), as in Remark 1 after Theorem 5.1 on p50 of [24].

Let \(( A, [\cdot, \cdot]_A) \) be a Lie algebra over \( k \), with finite dimension as a vector space over \( k \). If \( x \in A \), then we put \( \text{ad}_x = [x, \cdot]_A \) for every \( y \in A \), as usual. If \( \text{ad}_x \) is diagonalizable as a linear mapping from \( A \) into itself, then \( x \) is said to be ad-diagonalizable as an element of \( A \). Similarly, \( x \) is said to be ad-semisimple as an element of \( A \) when \( \text{ad}_x \) is semisimple as a linear mapping from \( A \) into itself, as on p24 of [13], and Definition 5.5 on p52 of [24].

Suppose that \( k \) is an algebraically closed field of characteristic 0, and that \( A \) is semisimple as a Lie algebra over \( k \). Let \( x \in A \) be given, so that \( \text{ad}_x \) is a linear mapping from \( A \) into itself. Remember that \( \text{ad}_x \) can be expressed in
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 \( \text{ad}_{x_1} \) and \( \text{ad}_{x_2} \), respectively, because the center of \( A \) is trivial, by semisimplicity. Similarly,

\[
(x_1, x_2)_A = 0,
\]

because \( \text{ad}_{x_1} \) and \( \text{ad}_{x_2} \) commute as linear mappings on \( A \). By construction, \( x_1 \) is ad-diagonalizable on \( A \), and \( x_2 \) is ad-nilpotent on \( A \). Note that \( x_1 \) 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 remarks on p24 of [13], and Theorem 5.6 on p52 of [24]. In particular, this is called the abstract Jordan decomposition of \( x \) in \( A \) in [13].

Let \( V \) be a finite-dimensional vector space over \( k \) again, and suppose that \( A \) is a Lie subalgebra of \( \text{gl}(V) \). 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 \( A = \text{sl}(V) \), as in [13]. Indeed, \( y_2 \) has trace 0 on \( V \), because \( y_2 \) is nilpotent on \( V \). This implies that \( y_1 \in \text{sl}(V) \) in this case, because \( x \in A \) by hypothesis. Some other situations will be discussed in the next section.

If \( w, z \in \text{gl}(V) \), then put

\[
(14.3.2) \quad \text{ad}_{\text{gl}(V), w}(z) = [w, z] = w \circ z - z \circ w,
\]

as usual. Let \( y \in \text{gl}(V) \) 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 \( y \in A \), and \( y_2 \) is nilpotent on \( V \), as in Section 10.8. W e have also seen that \( \text{ad}_{\text{gl}(V), y_1} \) is diagonalizable on \( \text{gl}(V) \), and \( \text{ad}_{\text{gl}(V), y_2} \) is nilpotent on \( \text{gl}(V) \), as in Section 10.9. Let \( A \) be a Lie subalgebra of \( \text{gl}(V) \). If \( w \in A \), then \( \text{ad}_{\text{gl}(V), w} \) maps \( A \) into itself, and the restriction of \( \text{ad}_{\text{gl}(V), w} \) 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 \( \text{ad}_{\text{gl}(V), y_1} \) commutes with \( \text{ad}_{\text{gl}(V), y_2} \) on \( \text{gl}(V) \). Under these conditions, we also have that

\[
(14.3.3) \quad \text{ad}_{A, y} = \text{ad}_{A, y_1} + \text{ad}_{A, y_2}.
\]

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 \( \text{ad}_{gl(V),x} \) maps \( A \) into itself, because \( x \in A \). Remember that \( \text{ad}_{\text{ad}_{gl(V),y_1}} \) and \( \text{ad}_{\text{ad}_{gl(V),y_2}} \) are the corresponding diagonalizable and nilpotent parts of \( \text{ad}_{gl(V),x} \), as a linear mapping from \( gl(V) \) into itself, as in Section 10.9. It follows that \( \text{ad}_{\text{ad}_{gl(V),y_1}} \) and \( \text{ad}_{\text{ad}_{gl(V),y_2}} \) map \( A \) into itself as well, as in Section 10.8. Let \( N = N_{gl(V)}(A) \) be the normalizer of \( A \) in \( gl(V) \), as in Section 9.8. The previous statement says that

\[
y_1, y_2 \in N.
\]

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 \( B_W \) 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

\[
A \subseteq B_W.
\]

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 \( B_W \) is a Lie subalgebra of \( gl(V) \). In particular, we can take \( W = V \), for which we get \( B_V = sl(V) \).

Let us check that

\[
y_1, y_2 \in B_W
\]

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 \( B_W \), over all \( A \)-submodules \( W \) of \( V \). Thus \( B \) is a Lie subalgebra of \( gl(V) \), and

\[
A \subseteq B,
\]
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

$$y_1, y_2 \in B,$$

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 7.14, 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 $T^p V$ 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

$$\prod_{j=1}^q \lambda_j(u_j)$$

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 \( T^q V' \) 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

\[
T^{p,q}V = T^p V \otimes T^q V',
\]

which may also be denoted \( V_{p,q} \), as on p40 of [24]. This reduces to \( T^q V' \) when \( p = 0 \), and to \( T^p V \) when \( q = 0 \). Note 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 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 \( \mu_j(v_l) = 1 \) 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

\[
(v_{j_1} \otimes \cdots \otimes v_{j_p}) \otimes (\mu_{l_1} \otimes \cdots \otimes \mu_{l_q}),
\]

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 \( a \in A \) and \( \lambda \in V' \), then put

\[
\rho^V_a (\lambda) = -\lambda \circ \rho^V_a,
\]

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 \( T^{p,q} V = T^q V' \) 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 (14.6.4) is a basis for $T^{p,q}V$, as in the previous section. It is easy to see that each of these elements (14.6.4) 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 $A$ as 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 $T^{p,q}V$ too, because $A_{p,q}$ is defined by adding and subtracting the $p+q$ linear mappings on $T^{p,q}V$, 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 $L(V)$ 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 $ad_A$ on $L(V)$, as in Section 10.9.
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

\[
\{ A \in gl(V) : A_{p,q}(t) = 0 \},
\]

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 \( T^{p,q}V \),

\[
A_{p,q} = (A_1)_{p,q} + (A_2)_{p,q},
\]

\((A_1)_{p,q} \) 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

\[
(A_1)_{p,q}(t) = (A_2)_{p,q}(t) = 0.
\]

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
\[(14.8.1) \quad T(W) \subseteq W.\]

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 \(T^{V/W}\) from \(V/W\) into itself such that
\[(14.8.2) \quad T^{V/W}(q(v)) = q(T(v)) \quad \text{for every } v \in V, \text{ as usual. It is easy to see that } T \mapsto T^{V/W} \text{ is an algebra homomorphism from } L_W(V) \text{ into } L(V/W). \text{ We shall normally be concerned with finite-dimensional vector spaces here, in which case the previous statement is more elementary.}\]

In particular, if \(T \in L_W(V)\) is nilpotent as a linear mapping on \(V\), then \(T^{V/W}\) is nilpotent as a linear mapping on \(V/W\). If \(R, T \in L_W(V)\) commute as linear mappings on \(V\), then \(R^{V/W}, T^{V/W}\) 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
\[(14.8.3) \quad T^{V/W}(q(v)) = q(T(v)) = q(\lambda v) = \lambda q(v), \quad \text{so that } q(v) \text{ is an eigenvector of } T^{V/W} \text{ with eigenvalue } \lambda \text{ as well. If } V \text{ is spanned by the eigenvectors of } T, \text{ then it follows that } V/W \text{ is spanned by the images under } q \text{ of the eigenvectors of } T, \text{ which are eigenvectors of } T^{V/W}.\]

Let \((A, [\cdot, \cdot]_A)\) and \((B, [\cdot, \cdot]_B)\) be Lie algebras over \(k\). Put
\[(14.8.4) \quad \text{ad}_{A,x}(y) = [x, y]_A, \quad \text{ad}_{B,w}(z) = [w, z]_B\]

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
\[(14.8.5) \quad \phi(\text{ad}_{A,x}(y)) = \phi([x, y]_A) = [\phi(x), \phi(y)]_B = \text{ad}_{B,\phi(x)}(\phi(y))\]

for every \(x, y \in A\). Suppose for the moment that \(\phi\) maps \(A\) onto \(B\). If \(x \in A\) and \(\text{ad}_{A,x}\) is nilpotent as a linear mapping from \(A\) into itself, then it follows that \(\text{ad}_{B,\phi(x)}\) 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 \(\text{ad}_{A,x}\), then \(\phi(y)\) is an eigenvector of \(\text{ad}_{B,\phi(x)}\), by (14.8.5). If \(A\) is spanned by the eigenvectors of \(\text{ad}_{A,x}\) and \(\phi\) maps \(A\) onto \(B\), then it follows that \(B\) is spanned by the eigenvectors of \(\text{ad}_{B,\phi(x)}\). 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 $V$, 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

\begin{equation}
\phi(w_1) = y_1, \quad \phi(w_2) = y_2,
\end{equation}

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 p7 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

\begin{equation}
ad_{A_2, w} = ad_{A_2, w'} + ad_{A_2, w''},
\end{equation}

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
to the abstract Jordan decomposition of $w$ in $A_2$, by uniqueness. This is the same as Exercise 9 on p31 of [13].

Let $B_1, B_2$ be finite-dimensional semisimple Lie algebras over $k$, and let $\psi$ be a Lie algebra homomorphism from $B_1$ into $B_2$. Theorem 5.7 on p52 of [24] basically says that $\psi$ maps abstract Jordan decompositions in $B_1$ to the corresponding abstract Jordan decompositions in $B_2$. Of course, this includes the case of subalgebras, as in the preceding paragraph. If $\psi$ maps $B_1$ onto $B_2$, 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 $\binom{n}{j} = n! / j! (n-j)!$ 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

\[
(T_1 + T_2)^n (T_1 + T_2) = \sum_{j=0}^{n} \binom{n}{j} T_1^{j+1} T_2^{n-j} + \sum_{j=0}^{n} \binom{n}{j} T_1^j T_2^{n-j+1}
\]

\[
= \sum_{j=1}^{n+1} \binom{n}{j-1} T_1^j T_2^{n+1-j} + \sum_{j=0}^{n} \binom{n}{j} T_1^j T_2^{n-j+1}.
\]

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 $A$ be an associative algebra over $k$ with a multiplicative identity element $e$. Suppose that $x \in A$ is nilpotent, so that $x^t = 0$ 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, \( x^0 \) 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
\[
\exp(x + y) = \sum_{n=0}^{\infty} \frac{1}{n!} (x + y)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{j=0}^{n} \binom{n}{j} x^j y^{n-j}
\]
(14.10.6)

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
\[
\exp(x) \exp(-x) = \exp(-x) (\exp x) = \exp 0 = e.
\]
(14.10.7)

This means that \( \exp x \) is invertible in \( A \), with
\[
(\exp x)^{-1} = \exp(-x).
\]
(14.10.8)

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 \( T^j(W) \subseteq W \) for every nonnegative integer \( j \), so that
\[
(\exp T)(W) \subseteq W.
\]
(14.10.9)

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 \( W \).
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

\[
(\exp \delta)(ab) = \sum_{n=0}^{\infty} (1/n!) \delta^n(ab) = \sum_{n=0}^{\infty} (1/n!) \sum_{j=0}^{n} \binom{n}{j} \delta^j(a) \delta^{n-j}(b)
\]

This is the same as the Cauchy product of the series defining \( (\exp \delta)(a) \) and \( (\exp \delta)(b) \), and the nilpotency condition on \( \delta \) ensures that all of these sums reduce to finite sums. Note that \( \exp \delta \) is invertible as a linear mapping from \( A \) into itself, as in the previous section.

More precisely, the double sum is the same as the Cauchy product of the series defining \( (\exp \delta)(a) \) and \( (\exp \delta)(b) \), and the nilpotency condition on \( \delta \) ensures that all of these sums reduce to finite sums. Note that \( \exp \delta \) is invertible as a linear mapping from \( A \) into itself, as in the previous section.

Thus (14.11.1) implies that \( \exp \delta \) is an algebra automorphism on \( A \), as on p9 of [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 \( \text{ad} \)-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

$$\phi \circ \text{ad}_x \circ \phi^{-1} = \text{ad}_{\phi(x)}$$

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

$$\phi \circ (\exp \text{ad}_x) \circ \phi^{-1} = \exp \text{ad}_{\phi(x)},$$

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

$$\exp \text{ad}_x(B) = B,$$

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 $\text{ad}_{B,x}$.

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 $a \cdot b$ 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 $L_a(x) = a \cdot x$ and $R_a(x) = x \cdot a$ for every $a \in A$. Note that

$$(L_a)^j = L_{a^j}, \quad (R_a)^j = R_{a^j}$$

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

$$\exp L_a = L_{\exp a}, \quad \exp R_a = R_{\exp a}$$

under these conditions.

Remember that $A$ may be considered as a Lie algebra over $k$ with respect to the corresponding commutator bracket $[a, b] = a \cdot b - b \cdot a$. If $a \in A$, then

$$\text{ad}_a = L_a - R_a = L_a + R_{-a},$$

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

$$\exp \text{ad}_a = \exp(L_a + R_{-a}) = (\exp L_a) \circ (\exp R_{-a})$$

$$= L_{\exp a} \circ R_{\exp(-a)} = L_{\exp a} \circ R_{(\exp a)^{-1}}.$$}

Equivalently,

$$\exp \text{ad}_a(x) = (\exp a) x (\exp a)^{-1}$$

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 $A$ and $B$ be associative algebras over $k$ with multiplicative identity elements $e_A$ and $e_B$, respectively. Suppose that $\phi$ is an algebra homomorphism from $A$ into $B$ such that $\phi(e_A) = e_B$. Let $x$ be a nilpotent element of $A$, which implies that $\phi(x)$ is nilpotent in $B$. Under these conditions, it is easy to see that

$$\phi(\exp_A x) = \exp_B \phi(x),$$

where $\exp_A x$ and $\exp_B \phi(x)$ are the exponentials of $x$ and $\phi(x)$ in $A$ and $B$, respectively.

Now let $(A, [\cdot, \cdot]_A)$ and $(B, [\cdot, \cdot]_B)$ be Lie algebras over $k$, and let $\psi$ be a Lie algebra homomorphism from $A$ into $B$. As usual, we put $\text{ad}_{A,x}(y) = [x, y]_A$ and $\text{ad}_{B,w}(z) = [w, z]_B$ for every $x, y \in A$ and $w, z \in B$. If $x, y \in A$, then

$$\psi(\text{ad}_{A,x}(y)) = \text{ad}_{B,\psi(x)}(\psi(y)),$$

as in Section 14.8. It follows that

$$\psi((\text{ad}_{A,x})^j(y)) = (\text{ad}_{B,\psi(x)})^j(\psi(y))$$

for every positive integer $j$. Suppose that $x$ is ad-nilpotent in $A$, and that $\psi(x)$ is ad-nilpotent in $B$. Thus $\exp \text{ad}_{A,x}$ and $\exp \text{ad}_{B,\psi(x)}$ can be defined as linear mappings from $A$ and $B$ into themselves, as before. Using (14.12.3), we get that

$$\psi((\exp \text{ad}_{A,x})(y)) = (\exp \text{ad}_{B,\psi(x)})(\psi(y))$$

for every $y \in A$.

Let $V$ be a module over $k$, and let $\rho$ be a representation of $A$ 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 $x \in A$ 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 $V$ unto 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 $y \in A$.

Suppose that $x$ is ad-nilpotent in $A$, in addition to $\rho_x$ being nilpotent on $V$. Thus $\exp \text{ad}_{A,x}$ is defined as a linear mapping from $A$ into itself, and

$$\rho(\exp \text{ad}_{A,x})(y) = (\exp \text{ad}_{\rho_x})(\rho_y)$$
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

$$\rho(\exp \text{ad}_{A,x})(y) = (\exp \rho_x) \circ \rho_y \circ (\exp \rho_x)^{-1}$$

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

$$\rho_x(v)_r = \sum_{j=1}^{m} \sum_{l=1}^{n} a_{j,l}^r x_j v_l,$$

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

$$\rho^K (v)_r = \sum_{j=1}^{m} \sum_{l=1}^{n} a_{j,l}^r x_j v_l$$

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

$$\rho_{[w,x]_{k^m}}(v) = \rho_w(\rho_x(v)) - \rho_x(\rho_w(v))$$
for every $w, x \in k^m$ and $v \in k^n$. This can be characterized by suitable conditions on the coefficients $a_{i,j}^r$ associated to $\rho$ and the structure constants for $[\cdot, \cdot]_{k^m}$. It follows that

$$(14.13.4) \quad \rho^K_{[w,x]_{k^m}}(v) = \rho^K_w(\rho^K_x(v)) - \rho^K(\rho^K_v(v))$$

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 $[\cdot,\cdot]_{K^m}$, 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 $T_K = \rho^K_{x_K}$. In particular, if $\rho^K_{x_K}$ is nilpotent as a mapping from $K^n$ into itself, then $\rho_x$ is nilpotent on $k^n$.

### 14.14 Lie’s theorem and nilpotency

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 $[A, A] \subseteq A$ is as defined in Section 9.2.

Suppose for the moment that $k$ is algebraically closed. Under these conditions, there is a flag $F = \{V_j\}_{j=0}^n$ in $V$ such that

$$(14.14.1) \quad \rho_a(V_j) \subseteq V_j$$
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

\[ \rho_a(V_j) \subseteq V_{j-1} \]  \hspace{1cm} (14.14.2)

for every $j = 1, \ldots, n$, as in Section 9.11. In particular, this implies that

\[ (\rho_a)^n = 0 \]  \hspace{1cm} (14.14.3)

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 $[,]_A$ 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, $[A_K, A_K]$ 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

\[ (\rho^K_a)^n = 0 \]  \hspace{1cm} (14.14.4)

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

$$(15.1.1) \quad x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$ 

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

$$(15.1.2) \quad [h, x] = 2 \cdot x, \quad [h, y] = -2 \cdot y, \quad \text{and} \quad [x, y] = h,$$

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

$$(15.1.3) \quad [\rho_h, \rho_x] = \rho_{[h,x]} = 2 \cdot \rho_x, \quad [\rho_h, \rho_y] = \rho_{[h,y]} = -2 \cdot \rho_y,$$

and $[\rho_x, \rho_y] = \rho_{[x,y]} = \rho_h$.

Equivalently, we may consider $V$ as a module over $sl_2(k)$, as a Lie algebra over $k$, with

$$(15.1.4) \quad a \cdot v = \rho_a(v)$$

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

$$(15.1.5) \quad V_\lambda = \{ v \in V : h \cdot v = \lambda v \},$$
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 $\langle h, x \rangle \leq v + x \leq \lambda v$. 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 $x \cdot v \in V_{\lambda + 2}$.

Similarly,

$$ h \cdot (y \cdot v) = ([h, y]) \cdot v + y \cdot (h \cdot v) = (-2 \cdot y) \cdot v + y \cdot (\lambda v) = (\lambda - 2) (y \cdot v), $$

where $\lambda - 2 = \lambda - 2 \cdot 1 = \lambda - 1 - 1$, as an element of $k$. This means that $y \cdot v \in V_{\lambda - 2}$.

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

$$ x \cdot v = 0, $$

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.
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 $\rho_h$ with eigenvalue $\lambda \in k$, then
\begin{equation}
\rho_h((\rho_x)^j(v)) = (\lambda + 2j)(\rho_x)^j(v)
\end{equation}
for every nonnegative integer $j$, as before. We also have that $(\rho_x)^j(v) = 0$ for some positive integer $j$, because $V$ has finite dimension. Let $j_0$ be the largest nonnegative integer such that $(\rho_x)^{j_0}(v) \neq 0$, so that $(\rho_x)^{j_0+1}(v) = 0$. This means that $(\rho_x)^{j_0}(v)$ 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
\begin{equation}
\{t \cdot v : t \in k\}
\end{equation}
is a one-dimensional linear subspace of $V$. If $v$ is primitive, then (15.2.3) is mapped into itself by $\rho_h$ and $\rho_x$. 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 $\rho_h$ and $\rho_x$. It follows that
\begin{equation}
2 \cdot \rho_x(v) = ([\rho_h, \rho_x])(v) = 0.
\end{equation}

If $k$ does not have characteristic 2, then we get that $\rho_x(v) = 0$, 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 $\rho_b$ 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
\begin{equation}
v_j = (1/j!)(\rho_y)^j(v)
\end{equation}
for every positive integer $j$, as on p32 of [13], and Theorem 3 on p18 of [23]. More precisely, $(\rho_y)^j$ is the $j$th power of $\rho_y$ 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.
15.3. RELATED SUBMODULES

Under these conditions, we have that

\( h \cdot v_j = (\lambda - 2j) v_j \)

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

\( y \cdot v_j = (j + 1) \cdot v_{j+1} \)

for every \( j \geq -1 \), by definition of \( v_j \). Let us check that

\( x \cdot v_j = (\lambda - j + 1) v_{j-1} \)

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

\( j x \cdot v_j = x \cdot (y \cdot v_{j-1}) = ([x,y]) \cdot v_{j-1} + y \cdot (x \cdot v_{j-1}). \)

Using induction, we may suppose that the analogue of (15.3.4) for \( j - 1 \) holds, so that

\( j x \cdot v_j = h \cdot v_{j-1} + (\lambda - (j-1) + 1) y \cdot v_{j-2}. \)

It follows that

\( j x \cdot v_j = (\lambda - 2 (j-1)) v_{j-1} + (\lambda - j + 2) ((j-2) + 1) v_{j-1}, \)

by (15.3.2) and (15.3.3). This reduces to

\( j x \cdot v_j = ((\lambda - 2j + 2) + (\lambda - j + 2) (j-1)) v_{j-1} \)

\( = (\lambda + (\lambda - j) (j - 1)) v_{j-1} = j (\lambda - j + 1) v_{j-1}. \)

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,

\( x \cdot v_j = (\lambda - j + 1) v_{j-1} \)

\( \text{for each } j \geq 0, \text{ we have that } v_j \neq 0. \)

This means that the \( v_j \)'s are linearly independent in \( V \), because they are eigenvectors for \( \rho_h \) 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,

\( v_l = 0 \text{ for some } l \geq 1. \)

This implies that \( v_l = 0 \) 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

\( x \cdot v_{m+1} = (\lambda - m) v_m, \)
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 \( \text{sl}_2(k) \), 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 \( \text{sl}_2(k) \), 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 \( \text{sl}_2(k) \). 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

\[
H(v_j) = (m - 2j) \cdot v_j,
\]

\[
Y(v_j) = (j + 1) \cdot v_{j+1},
\]

\[
X(v_j) = (m - j + 1) \cdot v_{j-1}
\]

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

\[
H(X(v_j)) - X(H(v_j)) = (m - j + 1) \cdot H(v_{j-1}) - (m - 2j) \cdot X(v_j)
\]

\[
= (m - j + 1)(m - 2(j - 1)) \cdot v_{j-1}
\]

\[
- (m - 2j) \cdot X(v_j)
\]

\[
= ((m - 2(j - 1)) - (m - 2j)) \cdot X(v_j) = 2 \cdot X(v_j)
\]
for each \(j = 0, 1, \ldots, m\). Similarly,

\[
H(Y(v_j)) - Y(H(v_j)) = (j + 1) \cdot H(v_{j+1}) - (m - 2j) \cdot Y(v_j)
\]

for every \(j = 0, 1, \ldots, m\). We also have that

\[
X(Y(v_j)) - Y(X(v_j)) = (j + 1) \cdot X(v_{j+1}) - (m - j + 1) \cdot Y(v_{j-1})
\]

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 \(\mathfrak{sl}_2(k)\), as in (15.1.1). Thus every element of \(\mathfrak{sl}_2(k)\) 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 \(\mathfrak{sl}_2(k)\) 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 \(a \in \mathfrak{sl}_2(k)\), 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 \(\mathfrak{sl}_2(k)\), 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 \(\mathfrak{sl}_2(k)\) on \(k\). If \(m = 1\), then \(\rho\) corresponds to the standard representation of \(\mathfrak{sl}_2(k)\) on \(k^2\). If \(m = 2\), then one can check that \(\rho\) is isomorphic to the adjoint representation on \(\mathfrak{sl}_2(k)\). 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 \(\mathfrak{sl}_2(k)\), 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 \(\mathfrak{sl}_2(k)\) 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 \( \rho^V \) 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 \( v_j \)’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

$$f(T) = f_1 T_1 + f_2 T_2,$$

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

$$x \cdot f(T) = f_2 T_1, \quad y \cdot f(T) = f_1 T_2, \quad \text{and} \quad h \cdot f(T) = f_1 T_1 - f_2 T_2.$$

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 $k[T_1, T_2]$ so that it is a derivation, as in Exercise 4 on p34 of [13]. Thus we would like to have that

$$a \cdot (f(T) g(T)) = (a \cdot f(T)) g(T) + f(T) (a \cdot g(T))$$

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

$$a \cdot (T_1^{\alpha_1} T_2^{\alpha_2}) = \alpha_1 \cdot (a \cdot T_1) T_1^{\alpha_1-1} T_2^{\alpha_2} + \alpha_2 \cdot (a \cdot T_2) T_1^{\alpha_1} T_2^{\alpha_2-1}.$$

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 $f(T), g(T) \in k[T_1, T_2]$. 

15.6 Another construction
Let \( \rho_a \) be the action of \( a \in sl_2(k) \) 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 \( a \in sl_2(k) \), 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 \( a \in sl_2(k) \) and \( d \geq 0 \), by construction. In order to show that this defines a representation of \( sl_2(k) \), as a Lie algebra over \( k \), on \( k[T_1, T_2] \), we should verify that
\[
[\rho_a, \rho_b] = \rho_{[a, b]}
\]
(15.6.5)
for every \( a, b \in sl_2(k) \). 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 \( sl_2(k) \) 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
\[
x \cdot (T_1^{\alpha_1} T_2^{\alpha_2}) = \alpha_2 \cdot T_1^{\alpha_1+1} T_2^{\alpha_2-1},
\]
(15.6.6)
\[
y \cdot (T_1^{\alpha_1} T_2^{\alpha_2}) = \alpha_1 \cdot T_1^{\alpha_1-1} T_2^{\alpha_2+1},
\]
(15.6.7)
\[
h \cdot (T_1^{\alpha_1} T_2^{\alpha_2}) = (\alpha_1 - \alpha_2) \cdot T_1^{\alpha_1} T_2^{\alpha_2}.
\]
(15.6.8)

If \( a \in sl_2(k) \), then \( \rho_a \) corresponds to the formal differential operator
\[
(a \cdot T_1) \partial_1 + (a \cdot T_2) \partial_2 = (a \cdot T_1) \frac{\partial}{\partial T_1} + (a \cdot T_2) \frac{\partial}{\partial T_2}
\]
(15.6.9)
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 \( y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \). 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
\[
\exp x = I + x = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \exp y = I + y = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},
\]
(15.7.1)
where \( I = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \) is the identity matrix in \( M_2(k) \). Similarly,
\[
\exp(-y) = I - y = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.
\]
(15.7.2)
Observe that
\[
(\exp x)(\exp -y) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix},
\]
(15.7.3)
and hence

\[(\exp x)(\exp -y)(\exp x) = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.\]

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 \(ad_x(z) = [x, z]\) and \(ad_y(z) = [y, z]\) for every \(z \in sl_2(k)\), as usual. Thus \(ad_x(x) = ad_y(y) = 0,\)

\[15.7.5\]

This implies that

\[(ad_x)^2(h) = (ad_y)^2(h) = 0.\]

Similarly,

\[15.7.7\]

It follows that \((ad_x)^3 = (ad_x)^3 = 0\) as mappings from \(sl_2(k)\) into itself, so that \(x\) and \(y\) are ad-nilpotent in \(sl_2(k)\).

This means that the exponentials of \(ad_x\) and \(ad_y\) are defined as linear mappings from \(sl_2(k)\) into itself, as in Section 14.10. In this situation, we have that

\[15.7.8\]

\[15.7.9\]

where \(I\) is the identity mapping on \(sl_2(k)\). Similarly,

\[15.7.10\]

Observe that

\[15.7.11\]

We also have that

\[15.7.12\]

Put

\[15.7.13\]

which defines a linear mapping from \(sl_2(k)\) into itself. More precisely, this is an inner automorphism of \(sl_2(k)\), as in Section 14.11. Observe that

\[15.7.14\]
Similarly,

\[ \sigma(y) = (\exp \text{ad}_x)((\exp -\text{ad}_y)(y + h - x)) \]
\[ = (\exp \text{ad}_x)(y + h - 2 \cdot y - x - h + y) = (\exp \text{ad}_x)(-x) = -x. \]

We also have that

\[ (15.7.15) \]
\[ \sigma(h) = (\exp \text{ad}_x)((\exp -\text{ad}_y)(h - 2 \cdot x)) \]
\[ = (\exp \text{ad}_x)(h - 2 \cdot y - 2 \cdot x - 2 \cdot h + 2 \cdot y) \]
\[ = (\exp \text{ad}_x)(-h - 2 \cdot x) = -h + 2 \cdot x - 2 \cdot x = -h. \]

It follows from (15.7.14), (15.7.15), and (15.7.16) that

\[ (15.7.17) \]
\[ \sigma(z) = -z^t \]
for every \( z \in \text{sl}_2(k) \), where \( z^t \) is the transpose of \( z \). This is the same as conjugating \( z \in \text{sl}_2(k) \) 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

\[ (15.8.1) \]
\[ \theta = (\exp X) \circ (\exp -Y) \circ (\exp X), \]
which is an invertible linear mapping from \( W(m) \) into itself.

Let \( \rho \) be the representation of \( \text{sl}_2(k) \) 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 \( \text{sl}_2(k) \). Remember that \( \exp \text{ad}_x \) and \( \exp \text{ad}_y \) are defined as invertible linear mappings on \( \text{sl}_2(k) \), because \( x \) and \( y \) are ad-nilpotent in \( \text{sl}_2(k) \).

If \( z \in \text{sl}_2(k) \), then

\[ (15.8.2) \]
\[ \rho_{(\exp \text{ad}_x)}(z) = (\exp \rho_x) \circ \rho_z \circ (\exp \rho_x)^{-1}, \]
as in Section 14.12. Similarly,

\[ (15.8.3) \]
\[ \rho_{(\exp -\text{ad}_y)}(z) = (\exp -\rho_y) \circ \rho_z \circ (\exp -\rho_y)^{-1} \]
for every \( z \in \text{sl}_2(k) \).

If \( \sigma \) is as in (15.7.13), then we get that

\[ (15.8.4) \]
\[ \rho_{\sigma(z)} = \theta \circ \rho_z \circ \theta^{-1} \]
for every $z \in \mathfrak{sl}_2(k)$. In particular,
\[
(15.8.5) \quad -Y = \rho_{\sigma(x)} = \theta \circ \rho_x \circ \theta^{-1} = \theta \circ X \circ \theta^{-1},
\]
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 p33-4 of [13], and Remark 1 on p21 of [23].

15.9 Additional properties of $sl_2(k)$ 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 $sl_2(k)$, as a Lie algebra over $k$, on $V = k^n$. This leads to a representation $\rho^{V_K}$ of $sl_2(K)$, 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 $k^n$ as 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 \( \text{sl}_2(k) \) module, then it follows that \( V \) is isomorphic as an \( \text{sl}_2(k) \) 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 \( V \) be a vector space over the real numbers. The complexification of \( V \) may be defined as a vector space over the complex numbers as follows. We start by taking \( V' = V \oplus V \) to be the direct sum of \( V \) with itself, as a vector space over the real numbers. Equivalently, \( V' \) is the Cartesian product \( V \times V \) of \( V \) with itself, where addition and scalar multiplication by real numbers is defined coordinatewise. If \( (v, w) \in V \times V \), then \( i (v, w) \) is defined as an element of \( V' \) by

\[
i (v, w) = (-w, v).
\]

(15.10.1)

It is easy to see that this makes \( V' \) into a vector space over the complex numbers. We shall normally identify \( v \in V \) with \( (v, 0) \in V \times V = V' \), so that \( V \) may be considered as a subset of \( V' \). Note that any basis for \( V \) as a vector space over the real numbers may be considered as a basis for \( V' \) 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,

\[
i V_0 = \{i v : v \in V_0\}
\]

(15.10.2)

is a real-linear subspace of \( V' \) as well. Note that \( V_0 + i V_0 \) 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

\[
V_0 \cap (i V_0) = \{0\}.
\]

(15.10.3)

This means that \( V_0 + i V_0 \) is isomorphic to the direct sum of \( V_0 \) and \( i V_0 \), as a vector space over the real numbers. Under these conditions, there is a natural isomorphism from the complexification of \( V_0 \) onto \( V_0 + i V_0 \), as vector spaces over the complex numbers, which is the identity mapping on \( V_0 \).

Let \( V' \) 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 \( V' \) 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 \( V' \) into \( W \) has a unique extension to a complex-linear mapping from the complexification \( V' \) of \( V \) into \( W \).

Let \( V' \) be a vector space over the real numbers, and let \( W' \) be the complexification of \( V' \). If \( \phi_1 \) is a linear mapping from \( V' \) into \( W' \), as a vector space over the real numbers, then \( \phi_1 \) may be considered as a real-linear mapping from \( V' \) into \( W' \).
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 $\mathbb{R}$. 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 $\mathbb{R}$. 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 $\mathbb{R}$, 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 $\mathbb{C}$. 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 $\mathbb{R}$ 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 $\mathbb{C}$, as before.

### 15.11 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 complex 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 \( \mathcal{L}(V_1, iW_1) \) 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 \( A_0 + i A_0 \), 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 \( A_0 + i A_0 \) is a subalgebra of \( A \), as an algebra over \( \mathbb{C} \) in the strict sense. In this case, \( A_0 + i A_0 \) 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
CHAPTER 15. COMPLEXIFICATIONS AND $SL_2(K)$ MODULES

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 $L(V_1)$ corresponds to a subalgebra of $L_C(V_2)$, as an algebra over $\mathbb{R}$. It follows that $L_C(V_2)$ can be identified with the complexification of $L(V_1)$ 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

\begin{equation}
\text{tr}_{V_1} \phi_1 = \text{tr}_{V_2} \phi_2, \tag{15.12.1}
\end{equation}

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

\begin{equation}
a \mapsto a^* \tag{15.13.1}
\end{equation}

be a conjugate-linear mapping from $A$ into itself. Put

\begin{equation}
A_{sa} = \{ a \in A : a^* = a \} \tag{15.13.2}
\end{equation}

and

\begin{equation}
A_{asa} = \{ a \in A : a^* = -a \}, \tag{15.13.3}
\end{equation}

which are the collections of vectors in $A$ that are self-adjoint and anti-self-adjoint with respect to (15.13.1), respectively. Clearly $A_{sa}$ and $A_{asa}$ are real-linear subspaces of $A$, with

\begin{equation}
A_{asa} = i A_{sa}. \tag{15.13.4}
\end{equation}
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 \(A_{\text{Lie}, C}\) be \(A\) considered as a Lie algebra over the complex numbers with respect to the corresponding commutator bracket, and let \(A_{\text{Lie}, R}\) be \(A_{\text{Lie}, C}\) considered as a Lie algebra over the real numbers. In this situation, \(A_{asa}\) is a Lie subalgebra of \(A_{\text{Lie}, R}\), and \(A_{\text{Lie}, C}\) 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 $s\ell(V)$, and in fact a Lie subalgebra of $s\ell(V)$, considered as a Lie algebra over the complex numbers. If $T \in s\ell(V)$, then $T^{*\beta} \in s\ell(V)$, and hence the self-adjoint and anti-self-adjoint parts of $T$ with respect to $\beta$ are contained in $s\ell(V)$ too. Thus $s\ell(V)$ 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 $\rho^V$, $\rho^W$ 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

$$\text{Hom}^A(V,W) = \text{Hom}_k^A(V,W)$$

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

$$\psi \circ \phi \circ \rho^V_a = \psi \circ (\rho^W_a \circ \phi) = \rho^Z_a \circ (\psi \circ \phi).$$

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 \( t I_V \) is an element of \( \text{Hom}_k^A(V, V) \).

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 \( L(V) \). Remember that \( L^\rho(V) \) may be used to denote the subalgebra of \( L(V) \) consisting of linear mappings 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 \( L^\rho(V) \) is invertible in \( L^\rho(V) \).

If \( k \) is also algebraically closed, and \( V \) has finite dimension as a vector space over \( k \), then \( L^\rho(V) \) 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

\[
\phi(\text{ad}_x(y)) = \text{ad}_x(\phi(y))
\]

for every \( x, y \in A \). This is the same as saying that

\[
\phi([x, y]_A) = [x, \phi(y)]_A
\]

for every \( x, y \in A \). If \( k \) is a field, then we may use \( L^{\text{ad}}(A) \) 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

\[
z = \sum_{j=1}^n \pi_j(z),
\]

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

\[
\pi_j([x, y]_A) = [x, \pi_j(y)]_A
\]
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
\[
[x, \phi(y)]_A = \phi([x, y]_A) = 0,
\]
using (16.1.5) in the first step. It follows that
\[
[x, \pi_j(\phi(y))]_A = 0,
\]
because \( [x, \pi_r(\phi(y))]_A = 0 \) automatically when \( r \neq j \). Equivalently, this means that \( \pi_j(\phi(y)) \) is an element of the center \( Z(A_j) \) of \( A_j \), as a Lie algebra over \( k \). If \( Z(A_j) = \{0\} \), then we get that
\[
\pi_j(\phi(y)) = 0.
\]
If \( Z(A_j) = \{0\} \) for every \( j = 1, \ldots, n \), then
\[
\phi(A_l) \subseteq A_l
\]
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 each \( j = 1, \ldots, n \), then we get that \( \mathcal{L}^{ad}(A) \) corresponds to the direct sum of \( \mathcal{L}^{ad}(A_l) \), \( 1 \leq l \leq n \). 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 \( \mathcal{L}^{ad}(A_j) \) 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 \( \mathcal{L}^{ad}(A) \) 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 $\lambda_i$’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

\begin{equation}
T_0(k^n) = k^{r_0}
\end{equation}

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 $k_1^n$, $k_1^r$, and $k_1^{r_0}$ 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 $k_1^n$ into $k_1^r$ that is linear over $k_1$, and let $\lambda_{l,k_1}$ be the natural extension of $\lambda_l$ to a linear functional on $k_1^n$ 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 $k_1^n$ into $k_1^{r_0}$ 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 $k_1^n$ is the same as the kernel of $T_{k_1}$, because of the property of $\lambda_{l,k_1}$ when $l > r_0$ mentioned in the preceding paragraph. It is easy to see that

\begin{equation}
T_{0,k_1}(k_1^n) = k_1^{r_0},
\end{equation}

using (16.3.1). This implies that $\lambda_{1,k_1}, \ldots, \lambda_{r_0,k_1}$ are linearly independent as linear functionals on $k_1^n$, and that the dimension of the kernel of $T_{0,k_1}$ in $k_1^n$ 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 $A$
16.5. SEMISIMPLICITY AND DIMENSIONS

Let $k$ be a field of characteristic 0, and let $(A, [\cdot, \cdot])$ be a Lie algebra over $k$ of positive finite dimension $n$. As in the previous section, we may as well take

\begin{equation}
([x, y])_r = \sum_{j=1}^{n} \sum_{l=1}^{n} c_{j,l}^{r} x_j y_l
\end{equation}

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 $c_{j,l}^{r}$ 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 $u_i$ is equal to 1 when $j = i$, and to 0 when $j \neq i$. 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

\begin{equation}
\phi([u_j, u_l]) = [u_j, \phi(u_l)]
\end{equation}

for every $j, l = 1, \ldots, n$. Of course, the space $\mathcal{L}(A)$ 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 $\mathcal{L}^{\text{ad}}(A)$ 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 $\mathcal{L}(A_{k_1})$ of linear mappings from $A_{k_1}$ into itself, as a vector space over $k_1$, corresponds to the space $M_n(k_1)$ of $n \times n$ matrices with entries in $k_1$. The space $\mathcal{L}^{\text{ad}}(A_{k_1})$ 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 $\mathcal{L}^{\text{ad}}(A_{k_1})$, as a vector space over $k_1$, is the same as the dimension of $\mathcal{L}^{\text{ad}}(A)$, 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
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^{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^{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^{ad}(A_{k_1}) is h, as a vector space over k_1. We also get that L^{ad}(A_{k_1}) is commutative as an algebra over k_1.

Similarly, the space L^{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^{ad}(A), as a vector space over k, is the same as the dimension of L^{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 φ is a linear mapping from A into itself, as a vector space over k, then φ has a natural extension to a linear mapping from A_{k_1} into itself, as a vector space over k_1. If φ is a homomorphism from A into itself, as a module over itself with respect to the adjoint representation, then the extension of φ 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^{ad}(A) is commutative as an algebra over k with respect to composition of mappings, because L^{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^{ad}(A) are invertible in L^{ad}(A), by Schur’s lemma, as in Section 16.1. This means that L^{ad}(A) is a field, because L^{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, . . . , A_m over k for some positive integer m, as in Section 10.15. This implies that L^{ad}(A) is isomorphic to the direct sum of L^{ad}(A_1), . . . , L^{ad}(A_m) as an associative algebra over k, as in Section 16.2. It follows that L^{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, [·, ·]) 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 $\mathcal{L}^{\text{ad}}(A_{k_1})$ 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, $\mathcal{L}^{\text{ad}}(A_{k_1})$ 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 $\mathcal{L}^{\text{ad}}(A_{k_1})$, as a vector space over $k_1$, is equal to one. It follows that the space $\mathcal{L}^{\text{ad}}(A)$ 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, $\mathcal{L}^{\text{ad}}(A)$ automatically contains all multiples of the identity mapping on $A$ by elements of $k$. This means that $\mathcal{L}^{\text{ad}}(A)$ consists exactly of multiples of the identity mapping on $A$ by elements of $k$ in this situation. This corresponds to the first part of part (c) of Exercise 5 on p55 of [24].

Conversely, suppose that $\mathcal{L}^{\text{ad}}(A)$ consists exactly of multiples of the identity mapping on $A$ by elements of $k$. This is the same as saying that $\mathcal{L}^{\text{ad}}(A)$ has dimension one as a vector space over $k$, as before. It follows that $\mathcal{L}^{\text{ad}}(A_{k_1})$ 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 $\mathcal{L}^{\text{ad}}(A_{k_1})$ 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 $K = \mathcal{L}^{\text{ad}}(A)$, 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, $\mathcal{L}^{\text{ad}}(A_K)$ 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.
Of course, $\mathcal{L}^{\text{ad}}(A_K)$ contains the multiples of the identity mapping on $A$ by elements of $K$, which correspond to linear mappings from $A$ into itself as a vector space over $k$. Conversely, if $\phi$ is any element of $\mathcal{L}^{\text{ad}}(A_K)$, then $\phi$ is linear as a mapping from $A$ into itself, as a vector space over $k$. This follows from the fact that $K$ contains the multiples of the identity mapping on $A$ by elements of $k$, by construction. We also have that $\phi$ is a homomorphism from $A$ into itself, as a module over itself as a Lie algebra over $k$, and with respect to the adjoint representation, because of the analogous property of $\phi$ as a mapping on $A_K$. This means that $\phi$ is an element of $\mathcal{L}^{\text{ad}}(A) = K$. Thus $\mathcal{L}^{\text{ad}}(A_K)$ consists exactly of multiples of the identity mapping on $A$ by elements of $K$. This implies that $A_K$ is absolutely simple as a Lie algebra over $K$, as before. This is the second part of part (c) of Exercise 5 on p55 of [24].

16.7 Algebras over subrings

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 a Lie algebra over $k_0$, then it follows that $A$ is semisimple as a Lie algebra over $k_0$.

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 $\text{tr}_{k/k_0} x$ 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
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

\[
\text{tr}_{k/k_0} x = n_0 \cdot x.
\]

Let \( n \) be a positive integer, so that \( V = k^n \) is a free module over \( k \) with respect to coordinatewise additional and scalar multiplication. We may also consider \( V \) as a free module over \( k_0 \) of rank \( n_0 \), 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 \( n \), 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

\[
\text{tr}_{V,k_0} T = \text{tr}_{k/k_0} (\text{tr}_{V,k} T).
\]

### 16.8 Lie algebras over subfields

Let \( k \) be a field, and let \( k_0 \) be a subfield of \( k \). Also let \((A, [\cdot, \cdot])\) 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

\[
b(x, y) = \text{tr}_A(\text{ad}_x \circ \text{ad}_y) \in k
\]

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

\[
b_0(x, y) = \text{tr}_{A_0}(\text{ad}_x \circ \text{ad}_y) \in k_0
\]

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

\[
b_0(x, y) = \text{tr}_{k/k_0} b(x, y),
\]

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

\[
b(x, y_1) = 1.
\]
This implies that
\[ b_0(x, y_1) = \text{tr}_{k/k_0} 1 = n_0 \cdot 1, \]
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
\[ \phi_t(x) = t x \]
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
\[ \phi_t(A_{0,j}) \subseteq A_{0,j} \]
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].
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

$$[y, z]_A = 0.$$  

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 $\text{ad}_{B, x} = 0$.  

331
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

\[
\text{ad}_{B,x}(y) = [x, y]_A = \lambda y.
\]

We would like to show that \( \lambda = 0 \).

Equivalently,

\[
\text{ad}_{B,y}(x) = -\lambda y,
\]

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 \( \text{ad}_{B,y} \), 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,

\[
\text{ad}_{B,y}(\text{ad}_{B,y}(x)) = -\lambda \text{ad}_{B,y}(y) = -\lambda [y, y]_A = 0,
\]

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

\[
\text{ad}_w(x) = [w, x]_A = \alpha(w) x
\]

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

\[
A_\alpha = \{ x \in A : \text{ for each } w \in B, \text{ ad}_w(x) = [w, x]_A = \alpha(w) x \}.
\]

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
\begin{equation}
A_0 = \{ x \in A : \text{for each } w \in B, \text{ ad}_w(x) = [w, x]_A = 0 \}.
\end{equation}
This is the same as the centralizer $C_A(B)$ of $B$ in $A$, as in Section 7.6. Note that
\begin{equation}
B \subseteq A_0 = C_A(B),
\end{equation}
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
\begin{equation}
\Phi_B = \{ \alpha \in B' : \alpha \neq 0 \text{ and } A_\alpha \neq \{0\} \}.
\end{equation}
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 ad$_u$ and ad$_v$ commute as linear mappings from $A$ into itself, as in Section 2.4. By hypothesis, ad$_u$ is diagonalizable as a linear mapping from $A$ into itself for each $u \in B$. It follows that the linear mappings 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 p45 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
\begin{equation}
[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.
\end{equation}
This uses the Jacobi identity in the first step, or, equivalently, the fact that ad$_w$ is a derivation on $A$. Thus
\begin{equation}
[x, y]_A \in A_{\alpha + \beta}
\end{equation}
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
\begin{equation}
\text{ad}_x(A_\gamma) \subseteq A_{\alpha + \gamma}
\end{equation}
for every $\gamma \in B'$, by (17.2.7). This implies that
\begin{equation}
(\text{ad}_x)^n(A_\gamma) \subseteq A_{n, \alpha + \gamma}
\end{equation}
for every positive integer $n$, where $(\text{ad}_x)^n$ is the $n$th power of 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
\begin{equation}
(\text{ad}_x \circ \text{ad}_y)(A_\gamma) = \text{ad}_x(\text{ad}_y(A_\gamma)) \subseteq A_{\alpha + \beta + \gamma}
\end{equation}
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
\[
 b([w, x], y) = -b(x, [w, y]) \tag{17.3.1}
\]
for every \( w, x, y \in A \), as in Sections 6.10 and 7.7. In particular, the Killing form
\[
 b_A(x, y) = \text{tr}_A(\text{ad}_x \circ \text{ad}_y) \tag{17.3.2}
\]
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 that
\[
 \alpha(w) b(x, y) = -\beta(w) b(x, y). \tag{17.3.3}
\]
If \( \alpha + \beta \neq 0 \), then there is a \( w \in B \) such that \( \alpha(w) \neq -\beta(w) \), and hence
\[
 b(x, y) = 0. \tag{17.3.4}
\]
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 \( y \in A \), then
\[
 b(w, [x, y]) = b([w, x], y) = \alpha(w) b(x, y). \tag{17.3.5}
\]
Similarly, if \( x \in A \) and \( y \in A_{-\alpha} \), then
\[
 b([x, y], w) = b(x, [y, w]) = \alpha(w) b(x, y). \tag{17.3.6}
\]
This corresponds to Theorem 3 (ii) on p44 of [23], and is related to part (c) of the proposition on p37 of [13]. Note that
\[
 [x, y] \in A_0 \tag{17.3.7}
\]
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
\[
 b(x, y) \neq 0, \tag{17.3.8}
\]
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
\[
 A_\alpha + A_{-\alpha} \tag{17.3.9}
\]
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

(17.4.1) \[ \text{tr}_V(R \circ T) = 0. \]

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,

(17.4.2) \[ C_A(B) = B, \]

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

(17.4.3) \[ C_A(B) \subseteq B. \]

Let $x \in C_A(B)$ be given, and let

(17.4.4) \[ x = x_1 + x_2 \]

be the abstract Jordan decomposition of $x$ in $A$, as in Section 14.3. Thus $x_1, x_2 \in A$,

(17.4.5) \[ \text{ad}_x = \text{ad}_{x_1} + \text{ad}_{x_2}, \]

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

(17.4.6) \[ x_1, x_2 \in C_A(B), \]

as in Step (1) of the proof on p36 of [13].

If $x \in C_A(B)$, then the linear span

(17.4.7) \[ B(x) = \{w + tx : w \in B, t \in k\} \]

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
This implies that \( B(x) = B \), because \( B \) is supposed to be maximal in \( A \), so that

(17.4.8) \( x \in B \).

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

(17.4.9) \( b_A(x, y) = 0 \)

for every \( y \in B \). 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

(17.4.10) \( b_A(x, z) = \text{tr}_A(\text{ad}_x \circ \text{ad}_z) = 0 \),

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 \( x_1 \in B \), 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

(17.4.11) \( \text{ad}_{C_A(B), x} = \text{ad}_{C_A(B), x_2} \)

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

(17.4.12) \( B \cap ([C_A(B), C_A(B)]) = \{0\} \).
To see this, observe that

\[(17.4.13) \quad b_{A}(x_1, x_2, y) = b_{A}(x_1, [x_2, y]_A) = 0\]

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

\[(17.4.14) \quad [C_A(B), C_A(B)] = \{0\} \neq 0\]

This implies that

\[(17.4.15) \quad ([C_A(B), C_A(B)]) \cap Z(C_A(B)) = \{0\},\]

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

\[(17.5.1) \quad B \subseteq C_A(B),\]
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

\[
B_1 \subseteq C_A(B_1) \subseteq C_A(B).
\]

(17.5.2)

If

\[
C_A(B) = B,
\]

(17.5.3)

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 \( B^\prime \) of \( B \), then we let \( A_\alpha \) be the set of \( x \in A \) such that

\[
\text{ad}_w(x) = [w, x]_A = \alpha(w) x
\]

(17.5.4)

for every \( w \in B \), as in Section 17.2. Let \( \Phi_B \) be the set of \( \alpha \in B^\prime \) 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

\[
\alpha(w) = b(w, t_{b, \alpha})
\]

(17.5.5)

for every \( w \in B \).

Let \( \alpha \in B' \), \( x \in A_\alpha \), and \( y \in A_{-\alpha} \) be given, so that

\[
[x, y]_A \in A_0 = C_A(B) = B.
\]

(17.5.6)

Remember that

\[
b(w, [x, y]_A) = \alpha(w) b(x, y)
\]

(17.5.7)
for every \( w \in B \), as in Section 17.3. It follows that

\[ [x, y]_A = b(x, y) t_{b, \alpha}, \quad (17.5.8) \]

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

\[ B_\alpha = [A_\alpha, A_{-\alpha}] \quad (17.5.9) \]

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 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

\[ \alpha(t_{b, \alpha}) = b(t_{b, \alpha}, t_{b, \alpha}) \neq 0 \quad (17.6.1) \]

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

\[ [t_{b, \alpha}, x]_A = \alpha(t_{b, \alpha}) x = 0 \quad (17.6.2) \]

for every \( x \in A_\alpha \), and that

\[ [t_{b, \alpha}, y]_A = -\alpha(t_{b, \alpha}) y = 0 \quad (17.6.3) \]

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

\[ [x, y]_A = t_{b, \alpha}, \quad (17.6.4) \]

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
\[
\alpha(h_\alpha) = 2\alpha(t_{b,\alpha})^{-1} \alpha(t_{b,\alpha}) = 2.
\]
(17.6.6)
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 \( y_\alpha \in A_{-\alpha} \) such that
\[
b(x_\alpha, y_\alpha) \neq 0.
\]
(17.6.7)
Using the facts that \( h_\alpha \in B \) and \( x_\alpha \in A_\alpha \) in the first step. Similarly,
\[
[h_\alpha, y_\alpha]_A = -\alpha(h_\alpha)y_\alpha = -2y_\alpha,
\]
(17.6.10)
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 \( (0 1 0), (0 0 1), \) and \( (1 -1 0) \) 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
\[
t_{b,-\alpha} = -t_{b,\alpha},
\]
(17.6.11)
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
\[
h_{-\alpha} = -h_\alpha,
\]
(17.6.12)
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,
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

$$\text{ad}_w(x) = [w, x]_A = \alpha(w) x$$

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

$$A_0 = C_A(B) = B,$$

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

$$\alpha(w) = b(w, t_{b,\alpha})$$

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

$$b(x_\alpha, y) = 0.$$  

It follows that

$$[x_\alpha, y]_A = 0,$$

by (17.5.8). Note that

$$[h_\alpha, y]_A = -\alpha(h_\alpha) y = -2y,$$

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

$$A_\alpha + A_{-\alpha} + B_\alpha$$

(17.7.7)

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, $y_\alpha \in A_{-\alpha}$ 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 $c \alpha \in \Phi_B$, (17.7.8)

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

$$[h_\alpha, z]_A = c \alpha(h_\alpha) z = 2cz,$$

(17.7.9)

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 $Q$ 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

$$2 \alpha \notin \Phi_B.$$  \hspace{1cm} (17.7.10)

If $\alpha/2$ were in $\Phi_B$, then we could apply the same argument to it, to get that $\alpha \notin \Phi_B$. Thus

$$\alpha/2 \notin \Phi_B.$$  \hspace{1cm} (17.7.11)

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 p39 of [13]. The second statement corresponds to the second part of Theorem 2 (a) on p43 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.
CHAPTER 17. SUBALGEBRAS AND DIAGONALIZABILITY

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
\[
\text{ad}_{h_\alpha}(y) = [h_\alpha, y]_A = \beta(h_\alpha) y.
\]
(17.8.1)

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
\[
z = (\text{ad}_{h_\alpha})^n(y) \quad \text{when } n \geq 0
\]
\[
= (\text{ad}_{h_\alpha})^{-n}(y) \quad \text{when } n \leq 0.
\]
(17.8.2)

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
\[
z \in A_{\beta - \beta(h_\alpha) \alpha},
\]
(17.8.3)

because \( y \in A_{\beta}, x_\alpha \in A_\alpha \), and \( y_\alpha \in A_{-\alpha} \), and using (17.2.7). In particular,
\[
A_{\beta - \beta(h_\alpha) \alpha} \neq \{0\},
\]
(17.8.4)

because \( z \neq 0 \).

Let us check that
\[
\beta - \beta(h_\alpha) \alpha \neq 0.
\]
(17.8.5)

Otherwise, if \( \beta - \beta(h_\alpha) \alpha = 0 \), then
\[
0 = \beta(h_\alpha) - \beta(h_\alpha) \alpha(h_\alpha) = \beta(h_\alpha) - 2\beta(h_\alpha) = -\beta(h_\alpha),
\]
(17.8.6)

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
\[
\beta - \beta(h_\alpha) \alpha \in \Phi_B,
\]
(17.8.7)

by (17.8.4). This corresponds to statement 2.9 on p46 of [23].

If \( \beta = \alpha \), then
\[
\beta - \beta(h_\alpha) \alpha = \alpha - \alpha(h_\alpha) \alpha = \alpha - 2\alpha = -\alpha,
\]
(17.8.8)

because \( \alpha(h_\alpha) = 2 \), as before. Similarly, if \( \beta = -\alpha \), then
\[
\beta - \beta(h_\alpha) \alpha = -\alpha - (\alpha(h_\alpha)) \alpha = -\alpha + 2\alpha = \alpha.
\]
(17.8.9)

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 $-\alpha \in \Phi_B$, 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

\[(17.8.10) \quad [h_\alpha, y]_A = 2 \alpha (h_\alpha) y = 4y.\]

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

\[(17.8.11) \quad \text{ad}_{x_\alpha}(y) = 0.\]

Observe that

\[(17.8.12) \quad \text{ad}_{h_\alpha}(y) = \text{ad}_{x_\alpha}(\text{ad}_{y_\alpha}(y)) - \text{ad}_{y_\alpha}(\text{ad}_{x_\alpha}(y)) = \text{ad}_{x_\alpha}(\text{ad}_{y_\alpha}(y)),\]

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

\[(17.8.13) \quad \text{ad}_{x_\alpha}(\text{ad}_{y_\alpha}(y)) = 0,\]

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

\[(17.9.1) \quad \beta \not= \pm \alpha.\]

Thus, for each $j \in \mathbb{Z}$,

\[(17.9.2) \quad \beta + j \alpha \not= 0,\]

as in the previous sections. Let $E$ be the linear span in $A$ of the subspaces

\[(17.9.3) \quad A_{\beta + j \alpha},\]

where $j \in \mathbb{Z}$. More precisely, we may restrict our attention to $j \in \mathbb{Z}$ such that

\[(17.9.4) \quad \beta + j \alpha \in \Phi_B,\]

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

\[
\beta(h_\alpha) + j \alpha(h_\alpha)
\]

(17.9.5)

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

\[
\beta(h_\alpha) + 2j,
\]

(17.9.6)

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

\[
m = \beta(h_\alpha) + 2q
\]

(17.9.7)

and

\[
-m = \beta(h_\alpha) - 2r,
\]

(17.9.8)

as in Section 15.3. In particular,

\[
\beta(h_\alpha) = r - q.
\]

(17.9.9)

Suppose that \( j \in \mathbb{Z} \) satisfies

\[
-r \leq j \leq q,
\]

(17.9.10)

so that

\[
\beta(h_\alpha) - 2r \leq \beta(h_\alpha) + 2j \leq \beta(h_\alpha) + 2q.
\]

(17.9.11)

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].
17.9. ADDING ELEMENTS OF $\Phi_B$

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

$$A_{\beta + (j+1) \alpha}.$$  

(17.9.12)  

This follows from Section 15.3 again. These properties of $E, q, r$ and $\beta$ 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

$$[A_\alpha, A_\beta] = A_{\alpha + \beta},$$

(17.9.14)

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.
Bibliography


Index

A((T)), 161
A[[T]], 80, 161
A[[T_1, \ldots, T_n]], 110
A[T], 80
A[T_1, \ldots, T_n], 110
A^op, 137
absolute convergence, 18
absolute value functions, 2
ultrametric, 3
absolute values, 179
absolutely simple Lie algebras, 327
absorbing sets, 21
abstract Jordan decompositions, 286
ad-diagonalizable elements, 285
ad-nilpotent elements, 190
ad-semisimple elements, 285
adjoint mappings, 17, 45, 47, 64, 66, 231, 248
algebra automorphisms, 27, 33, 34
algebra homomorphisms, 27, 33, 34
algebra isomorphisms, 27, 33, 34
algebras, 26
anti-self-adjointness, 33, 72, 318
linear mappings, 46, 247
antisymmetric mappings, bilinear, 26, 42
linear, 42
multilinear, 144
archimedean absolute value functions, 3
archimedean property, 180
associative algebras, 27
associative bilinear mappings, 135
balanced sets, 21
balanced subgroups, 23
Banach algebras, 57
Banach spaces, 11
bilinear actions, 122
bilinear forms, 42
bilinear mappings, 19, 26
BL(V), 51
BL(V, W), 12
BL(V_1, \ldots, V_n; W), 145
bounded bilinear mappings, 19
bounded linear mappings, 12
bounded multilinear mappings, 145
C, 2
c(X, W), 158
c_00(X, W), 158
C_A(E), 148
Casimir elements, 271, 273, 274
Cauchy products, 78
Cauchy–Schwarz inequality, 16
center, 148, 149
centralizer, 148
closed balls, 2
commutative algebras, 27
commutative Lie algebras, 29
complete reducibility, 269
complex-linear mappings, 15
complex-linear subspaces, 314
complexifications, 314
conjugate-linear mappings, 15, 246
conjugation in \mathbb{k}[\sqrt{-a}], 245
convergent series, 17
D_n(A), 206
d_n(A), 206
derivations, 31
derivatives, 106, 114
derived algebra, 184
derived series, 184
descending central series, 187
diagonal matrices, 206
diagonalizable linear mappings, 210
direct products, 76
direct sums, 76
discrete absolute value functions, 6
discrete metric, 1
equivalent absolute value functions, 5
extended formal power series, 180
flags, 194
formal differential operators, 116, 120
formal polynomials, 80, 110, 119
formal power series, 80, 110
general linear algebra, 38, 39
general linear group, 50
$GL(V)$, 50
$gl(V)$, 39
$GL_n(A)$, 50
$gl_n(A)$, 38
Hermitian forms, 46, 247
Hilbert spaces, 16
$\text{Hom}(A, B)$, 26
$\text{Hom}_k(A, B)$, 26
homomorphisms
between algebras, 27, 33, 34
between Lie algebras, 29
between modules, 25, 130, 131
between representations, 130, 131
ideals, 40, 41
identity matrix, 37
inner automorphism, 297
inner derivations, 282
inner products, 15, 234, 248
$\text{Int} A$, 296
intertwining mappings, 124, 130, 131
invariant bilinear mappings, 63
invariant elements, 149
invariant sesquilinear forms, 66
invertible elements, 49
invertible linear mappings, 59
involutions, 33, 34, 319
irreducible representations, 140
isometric linear mappings, 59
isomorphic ordered fields, 181
isomorphic representations, 130, 131
Jacobi identity, 28
kernels of module homomorphisms, 40
killing form, 153
$\mathcal{L}(V)$, 39
$\mathcal{L}(V, W)$, 39
$L(V_1, \ldots, V_n; W)$, 143
$L_k(V_1, \ldots, V_n; W)$, 143
Lie algebra homomorphisms, 29
Lie algebras, 29
Lie bracket, 29
Lie subalgebras, 29
Lie’s theorem, 190
linear mappings, 26
lower central series, 187
$M_n(A)$, 36
maximal vector of weight $\lambda$, 303
metrics, 1
module automorphisms, 63
module homomorphisms, 25, 130, 131
module isomorphisms, 26, 130, 131
modules,
over associative algebras, 127
over $k$, 25
over Lie algebras, 128
multi-indices, 109
multilinear mappings, 143
multiplicative identity elements, 27
multiplicative seminorms, 51
nilpotent elements of associative algebras, 190
nilpotent Lie algebras, 187
non-archimedean absolute value functions, 3
nondegenerate bilinear forms, 44
nondegenerate sesquilinear forms, 47, 247
nonnegative self-adjoint operators, 263
normalizer, 191
norms, 10
open balls, 1
open sets, 2
opposite algebra, 137
ordered fields, 178
ordered rings, 178
orthogonal transformations, 61
orthogonal vectors, 234, 248
Ostrowski’s theorems, 5, 6

\( p \)-adic absolute value, 3
\( p \)-adic integers, 8
\( p \)-adic metric, 3
\( p \)-adic numbers, 5
poles of bounded order, 162, 164, 167, 168, 170
positive formal series, 180
positive self-adjoint operators, 263
primitive of weight \( \lambda \), 303
products of submodules, 183

\( \mathbb{Q} \), 3
\( \mathbb{Q}_p \), 5

\( \mathbb{R} \), 2
\( \mathbb{R}_+ \), 6
radicals of bilinear mappings, 155
radicals of Lie algebras, 186
real parts, 255
real-bilinear mappings, 315
real-linear subspaces, 314
real-linear mappings, 15, 314
reductive Lie algebras, 227
replicas, 217
representations,
of associative algebras, 126
of Lie algebras, 128
residue fields, 9

\( S_n(A) \), 206
\( s_n(A) \), 206
scalar matrices, 206
Schur’s lemma, 140
self-adjoint linear mappings, 46, 247
self-adjoint sets, 237, 251
self-adjointness, 33, 72, 318
self-normalizing Lie subalgebras, 192

semi-ultrametrics, 1
semi-ultranorms, 1
semimetrics, 1
seminorms, 10
semisimple Lie algebras, 186
semisimple linear mappings, 285
semisimple modules, 269
sesquilinear forms, 46, 247
simple Lie algebras, 222
simple modules, 140
\( SL(V) \), 50
\( sl(V) \), 40
\( SL_n(A) \), 50
\( sl_n(A) \), 38
solvable Lie algebras, 184
special linear algebra, 38, 40
special linear group, 50
subalgebras, 27, 29
submodules, 25, 129, 130
submultiplicative seminorms, 51
subrepresentations, 129
supports of functions, 158
symmetric mappings,
  bilinear, 26, 42
  linear, 42
  multilinear, 144
symmetric tensors, 160

\( T_{n,r}(A) \), 196
\( t_{n,r}(A) \), 196
tensor products, 157
tensor powers, 160
termwise eventual agreement, 81, 162
termwise eventually constant, 81, 162
toral subalgebras, 331
totally real linear subspaces, 314
trace, 38, 40, 152
transpose matrices, 37
trivial absolute value function, 2
trivial ultranorm, 11

\( \mathcal{U}_b(\mathcal{F}) \), 194
ultrametric absolute values, 3
ultrametrics, 1
ultranorms, 10
unitary transformations, 61
vanishing to some order, 81, 119
weight spaces, 303
weight $\alpha$, 332
weight $\lambda$, 303
weights, 303
Weyl’s theorem, 269, 276

$\mathbb{Z}$, 8
$\mathbb{Z}(A)$, 149
$\mathbb{Z}_+$, 3
$\mathbb{Z}_p$, 8