

Some basic topics in analysis

Stephen Semmes
Rice University

Preface

These informal notes deal with some possible topics for a second course in analysis. In particular, the reader is assumed to be familiar with metric spaces, sequences and series, and continuous functions. Some topics may be mentioned in a first course, with some review or elaboration here.

Of course, there are many textbooks in analysis, a few of which are mentioned in the bibliography. The aim here is to complement these textbooks, while trying to look ahead a bit to more advanced courses.

Although some basic notions are used frequently throughout the text, there is a fair amount of independence between the various sections and chapters. Thus the reader may wish to focus more on some parts, at least initially.

Some aspects of history related to topics like those considered here may be found in [4, 24, 27, 28, 29, 30, 42, 43, 62, 63, 67, 85, 94, 95, 96, 99, 101, 102, 103, 104, 105, 106, 108, 109, 120, 135, 153, 205, 224], for instance. Some songs related to some topics like those considered here may be found in [159, 160, 161]. Some remarks concerning the clarity of explanations in mathematics may be found in [107]. The reader may also be interested in [22].

Contents

1	Some mappings, metrics, and norms	1
1.1	Lipschitz mappings	1
1.1.1	A simple relationship between metrics	1
1.1.2	Square roots of metrics	1
1.2	Lipschitz conditions on \mathbf{R}	2
1.2.1	The definition of the derivative	2
1.2.2	Using the mean value theorem	2
1.3	Norms on \mathbf{R}^n	3
1.3.1	The definition of a norm	3
1.3.2	Some basic examples of norms	4
1.3.3	Metrics associated to norms	4
1.3.4	The Cauchy-Schwarz inequality for finite sums	5
1.4	Norms on \mathbf{C}^n	5
1.4.1	A larger family of norms	7
1.5	Some basic inequalities	7
1.5.1	Inequalities in the other direction	7
1.5.2	Some consequences of these inequalities	8
1.5.3	Estimates for arbitrary norms	8
1.6	Functions with finite support	9
1.6.1	Norms on $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$	9
1.6.2	Some inequalities between these norms	11
1.6.3	Cauchy sequences in $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$	11
1.7	Cauchy sequences and completeness	12
1.7.1	Cauchy sequences in subsets of metric spaces	12
1.7.2	Completeness of subsets	13
1.7.3	Cauchy sequences and uniform continuity	13
1.8	Pointwise and uniform convergence	13
1.8.1	An example on $[0, 1]$	13
1.8.2	Uniform convergence and continuity	14
1.8.3	Uniform convergence and uniform continuity	15
1.8.4	Some limits of limits	15
1.9	Bounded sets	15
1.9.1	Boundedness of compact sets	16
1.9.2	More on bounded sets	16

1.9.3	Closed balls	17
1.9.4	Bounded sequences	17
1.10	Some remarks and examples	18
1.10.1	A basis for $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$	18
1.10.2	Some more examples in $c_{00}(\mathbf{Z}_+, \mathbf{R})$	19
1.11	Bounded functions	19
1.11.1	Uniform convergence of bounded functions	20
1.11.2	The supremum metric	20
1.11.3	Convergence with respect to the supremum metric	21
1.12	Completeness of $\mathcal{B}(X, Y)$	21
1.12.1	The proof of completeness	21
1.12.2	Necessity of completeness of Y	22
1.13	More on bounded functions	22
1.13.1	The supremum norm	23
1.14	Continuous functions	23
1.14.1	The supremum metric and $C_b(X, Y)$	24
1.14.2	Completeness of $C_b(X, Y)$	24
1.14.3	Real and complex-valued functions	24
1.15	Continuous functions on $[0, 1]$	25
1.15.1	Norms on $C([0, 1], \mathbf{R})$, $C([0, 1], \mathbf{C})$	25
1.15.2	The L^1 norm on $C([0, 1], \mathbf{R})$ and $C([0, 1], \mathbf{C})$	25
1.15.3	The L^2 norm on $C([0, 1], \mathbf{R})$ and $C([0, 1], \mathbf{C})$	26
1.15.4	More on these integral norms	26
2	Basic ℓ^1 and ℓ^2 spaces	28
2.1	Infinite series	28
2.1.1	Series with nonnegative terms	29
2.1.2	Absolute convergence	29
2.2	Basic ℓ^1 spaces	30
2.2.1	More on $\ell^1(\mathbf{Z}_+, \mathbf{R})$, $\ell^1(\mathbf{Z}_+, \mathbf{C})$	31
2.3	Basic ℓ^2 spaces	32
2.3.1	The Cauchy-Schwarz inequality for infinite series	32
2.3.2	The triangle inequality for the ℓ^2 norm	33
2.3.3	More on $\ell^2(\mathbf{Z}_+, \mathbf{R})$, $\ell^2(\mathbf{Z}_+, \mathbf{C})$	34
2.4	Completeness of ℓ^1 , ℓ^2	34
2.4.1	Completeness of ℓ^1	35
2.4.2	Completeness of ℓ^2	36
2.5	Vanishing at infinity	36
2.5.1	Sequences converging to 0	37
2.5.2	More on vanishing at infinity	37
2.5.3	Closed sets in $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$	37

3	Some more metric spaces	39
3.1	Continuous functions on closed intervals	39
3.1.1	Norms on $C([a, b], \mathbf{R})$, $C([a, b], \mathbf{C})$	39
3.1.2	The L^1 norm and metric on $C([a, b], \mathbf{R})$, $C([a, b], \mathbf{C})$	40
3.2	The square norm	40
3.2.1	The Cauchy–Schwarz inequality for integrals	41
3.2.2	The triangle inequality for the L^2 norm	41
3.2.3	More on the L^2 norm	42
3.3	Riemann–Stieltjes integrals	42
3.3.1	Some basic properties of Riemann–Stieltjes integrals	43
3.3.2	Riemann–Stieltjes integrability	43
3.4	Riemann–Stieltjes integrals and seminorms	44
3.4.1	Some integral seminorms	44
3.4.2	More on these integral seminorms	45
3.4.3	Connections with uniform convergence	45
3.5	Square seminorms	46
3.5.1	Another Cauchy–Schwarz inequality for integrals	46
3.5.2	The triangle inequality for these square seminorms	47
3.5.3	More on these square seminorms	47
3.6	Some more monotone functions	48
3.6.1	Some homogeneity properties	48
3.7	Continuous functions on \mathbf{R}	49
3.7.1	Norms on $C_{com}(\mathbf{R}, \mathbf{R})$, $C_{com}(\mathbf{R}, \mathbf{C})$	49
3.7.2	Some integrals on \mathbf{R}	50
3.7.3	Some integral norms on $C_{com}(\mathbf{R}, \mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$	51
3.8	Riemann–Stieltjes integrals on \mathbf{R}	51
3.8.1	Some more integral seminorms	52
3.8.2	Some additional monotone functions	52
3.8.3	Another version of the Cauchy–Schwarz inequality	53
3.8.4	The case of bounded α	54
3.9	Dense sets and Weierstrass’ theorem	54
3.9.1	Density in $C_b(X, Y)$	55
3.9.2	Weierstrass’ theorem	55
3.10	Functions on sets	56
3.10.1	Linear subspaces of $c(X, \mathbf{R})$, $c(X, \mathbf{C})$	56
3.10.2	Subalgebras of $c(X, \mathbf{R})$, $c(X, \mathbf{C})$	56
3.10.3	Invariance under complex conjugation	57
3.11	More on collections of functions	57
3.11.1	Some continuous real-valued functions	57
3.11.2	More on subalgebras of $c(X, \mathbf{R})$, $c(X, \mathbf{C})$	58
3.11.3	Subalgebras and polynomials	58
3.12	The Stone–Weierstrass theorem	58
3.12.1	Some remarks about Weierstrass’ theorem	59
3.12.2	The complex version of the Stone–Weierstrass theorem	59
3.13	Algebras of bounded functions	60
3.13.1	Some remarks about closures	60

3.13.2	Some more remarks about closures	61
3.14	Some remarks about closed subalgebras	62
3.14.1	Closed subalgebras of $\ell^\infty(X, \mathbf{R})$	62
3.14.2	Closed subalgebras of $C(X, \mathbf{R})$	63
3.14.3	Using the g_x 's	64
4	Compactness and completeness	65
4.1	Diameters of sets	65
4.1.1	Some properties of the diameter	65
4.1.2	Some nested intersections	66
4.1.3	Necessity of completeness	67
4.2	Totally bounded sets	67
4.2.1	Using closed balls	68
4.2.2	Using sets with small diameter	68
4.2.3	Another characterization of total boundedness	69
4.3	Separable metric spaces	69
4.3.1	A characterization of separability	70
4.3.2	The definition of a base for the topology of X	70
4.3.3	Separability and countable bases	70
4.4	Lindelöf's theorem	71
4.4.1	The proof of Lindelöf's theorem	71
4.4.2	The Lindelöf property	72
4.5	The limit point property	72
4.5.1	Using compactness to get the limit point property	73
4.5.2	The limit point property implies total boundedness	73
4.5.3	Some more remarks about the limit point property	73
4.6	Countable open coverings	74
4.6.1	Using the limit point property to get compactness	75
4.6.2	Subsets of X with the limit point property	75
4.7	Sequential compactness	75
4.7.1	The limit point property and sequential compactness	75
4.7.2	The converse	76
4.7.3	More on sequential compactness	76
4.8	The Cauchy subsequence property	77
4.8.1	The small subsequence property	77
4.8.2	Total boundedness and small subsequences	77
4.8.3	Small subsequences and Cauchy subsequences	78
4.9	A criterion for compactness	78
4.9.1	Using the Cauchy subsequence property	78
4.9.2	A reformulation of total boundedness	79
4.9.3	Another proof of compactness	79
4.10	Cauchy sequences and convergent subsequences	80
4.10.1	Some additional remarks about compactness	80
4.11	The Baire category theorem	81
4.11.1	Some infinite intersections	81
4.11.2	The beginning of the proof	81

4.11.3	Repeating the process	82
4.11.4	The rest of the proof	82
4.12	The interior of a set	83
4.12.1	Some finite unions of closed sets in X	84
4.12.2	A reformulation of the Baire category theorem	84
4.12.3	First and second category	84
5	Equicontinuity and sequences of functions	86
5.1	Pointwise convergent subsequences	86
5.1.1	Complex and \mathbf{R}^n -valued functions	87
5.1.2	Finite sets E	87
5.1.3	Countably infinite sets E	87
5.1.4	A single subsequence	88
5.2	Equicontinuity	88
5.2.1	Uniform equicontinuity	89
5.2.2	Equicontinuity and sequences of mappings	89
5.2.3	Uniformly equicontinuous sequences of mappings	90
5.2.4	Lipschitz conditions and pointwise convergence	90
5.3	Uniformly Cauchy sequences	90
5.3.1	Uniform convergence implies uniformly Cauchy	91
5.3.2	Convergence when Y is complete	91
5.3.3	Equicontinuity of uniformly Cauchy sequences	91
5.4	Equicontinuity and uniform convergence	92
5.4.1	A localized uniform convergence property	92
5.4.2	Uniform convergence on compact sets	93
5.4.3	Another uniform convergence property	93
5.5	Equicontinuity and Cauchy sequences	94
5.5.1	Some initial steps	94
5.5.2	Using the Cauchy condition at a nearby point	94
5.6	Equicontinuity and subsequences	95
5.6.1	Pointwise convergence on X	95
5.6.2	The limit f of $\{f_{j_i}\}_{i=1}^\infty$	95
5.6.3	Some additional remarks	96
5.7	Pointwise and uniform boundedness	96
5.7.1	Pointwise boundedness and equicontinuity	97
5.7.2	Equicontinuity on X	97
5.8	Total boundedness in $\mathcal{B}(X, Y)$	98
5.8.1	Equicontinuity using total boundedness	99
5.8.2	Some more total boundedness conditions	100
5.9	Total boundedness using equicontinuity	100
5.9.1	Using uniform equicontinuity	100
5.9.2	Total boundedness of \mathcal{E}	101
5.10	Equiconvergence of limits	102
5.10.1	Equiconvergence and equicontinuity	102
5.10.2	Equiconvergence for sequences of mappings	102
5.10.3	Pointwise convergent sequences	103

5.10.4	Equiconvergence and uniform Cauchy sequences	103
5.11	Equiconvergence and differentiability	103
5.11.1	A criterion for equiconvergence	103
5.11.2	Sequences of differentiable mappings	104
5.11.3	Pointwise convergence and differentiability	104
5.11.4	Uniformly Cauchy difference quotients	104
6	More on sums and norms	106
6.1	Weierstrass' criterion	106
6.1.1	Continuity of the sum	107
6.1.2	Power series	107
6.1.3	Another continuity property	108
6.2	Radius of convergence	108
6.2.1	Defining the radius of convergence	108
6.2.2	More on R	109
6.2.3	Another convergence property	109
6.3	Termwise differentiation	110
6.3.1	Continuously-differentiable functions on $[a, b]$	111
6.3.2	Differentiating power series	111
6.3.3	More on differentiating power series	112
6.4	Rearrangements	112
6.4.1	Rearrangements with nonnegative terms	113
6.4.2	Rearrangements and absolute convergence	113
6.5	Cauchy products	113
6.5.1	Nonnegative a_j, b_l	114
6.5.2	Absolutely convergent series	115
6.5.3	Another type of argument	115
6.6	The complex exponential function	116
6.6.1	Exponentials of sums	116
6.6.2	Some more properties of $E(z)$	117
6.6.3	Exponentials and complex conjugation	117
6.6.4	Differentiating $\exp(at)$	118
6.7	Abel summability	118
6.7.1	The proof of Abel's theorem	119
6.7.2	Estimating $A(r) - s$	119
6.7.3	The limit as $r \rightarrow 1-$	120
6.8	Abel summability and Cauchy products	120
6.8.1	Abel summability conditions	121
6.8.2	A corollary about Cauchy products	121
6.9	Cesaro means	121
6.9.1	Cesaro summability of infinite series	121
6.10	Comparing norms on $\mathbf{R}^n, \mathbf{C}^n$	122
6.10.1	Using the standard basis	122
6.10.2	Comparison with the norm $\ \cdot\ _1$	123
6.10.3	Comparison with the norm $\ \cdot\ _2$	123
6.10.4	Comparison with the norm $\ \cdot\ _\infty$	124

6.11	Another comparison	124
6.11.1	Continuity of N	125
6.11.2	Using the extreme value theorem	125
6.11.3	Another comparison with $\ \cdot\ _2$	125
6.12	The standard inner products on $\mathbf{R}^n, \mathbf{C}^n$	126
6.12.1	Using complex numbers	126
6.12.2	The standard (Hermitian) inner product on \mathbf{C}^n	126
6.13	Sums and inner products	127
6.13.1	Another Hermitian inner product	127
6.13.2	Inner products on $\ell^2(\mathbf{Z}_+, \mathbf{R}), \ell^2(\mathbf{Z}_+, \mathbf{C})$	128
6.14	Some integral inner products	129
6.14.1	Some inner products on $C([a, b], \mathbf{C})$	129
6.14.2	Some inner products on $C_{com}(\mathbf{R}, \mathbf{R})$	130
6.14.3	Some inner products on $C_{com}(\mathbf{R}, \mathbf{C})$	131
7	Matrix norms and Lipschitz conditions	132
7.1	Real and complex matrices	132
7.1.1	Norms on matrices	132
7.1.2	The Hilbert–Schmidt norm	133
7.2	Matrices and linear mappings	133
7.2.1	Linear mappings associated to matrices	134
7.2.2	Matrices associated to linear mappings	134
7.2.3	Another description of the Hilbert–Schmidt norm	134
7.2.4	A simple estimate using the Hilbert–Schmidt norm	135
7.3	Some related estimates	135
7.3.1	Estimates in terms of $\ \cdot\ _1, \ \cdot\ _2$, and $\ \cdot\ _\infty$	135
7.3.2	Using an arbitrary norm on \mathbf{R}^n or \mathbf{C}^n	136
7.4	Spaces of linear mappings	137
7.4.1	Norms on these spaces of linear mappings	137
7.5	Operator norms	138
7.5.1	Two other characterizations of the operator norm	138
7.5.2	Checking that the operator norm is a norm	139
7.5.3	Comparing this with the Hilbert–Schmidt norm	140
7.5.4	The $n = m$ case	140
7.6	Determinants and volumes	141
7.6.1	Linear mappings and volumes	141
7.6.2	Determinants and operator norms	142
7.6.3	Volumes and operator norms	142
7.6.4	Eigenvalues and operator norms	143
7.7	Lipschitz constants	143
7.7.1	The definition of $\text{Lip}(f)$	143
7.7.2	Operator norms and Lipschitz constants	144
7.7.3	Vector-valued Lipschitz functions	144
7.8	Compositions and isometries	145
7.8.1	Isometric mappings	145
7.8.2	Composing linear mappings	146

7.8.3	Compositions and operator norms	146
7.8.4	Isometric linear mappings	146
7.9	Bilipschitz embeddings	147
7.9.1	Bilipschitz mappings	147
7.9.2	Compositions of bilipschitz mappings	148
7.9.3	A condition on linear mappings	148
7.9.4	More on composing linear mappings	148
7.10	Linear mappings and seminorms	149
7.10.1	Using a norm N	149
7.10.2	The $m = n$ case	149
7.11	Small perturbations	150
7.11.1	Analogous arguments for linear mappings	151
7.12	The contraction mapping theorem	151
7.12.1	Existence of the fixed point	152
7.12.2	Brouwer's fixed-point theorem	153
7.13	Norms on \mathbf{R}^n and completeness	153
7.13.1	Small perturbations of the identity mapping on \mathbf{R}^n	153
7.14	A localized condition	154
7.14.1	Completeness of $\overline{B}_N(r)$	154
7.14.2	Lipschitz mappings on $\overline{B}_N(r)$	155
7.14.3	Using the contraction mapping theorem on $\overline{B}_N(r)$	155
7.15	Open mappings	156
7.15.1	A helpful criterion on \mathbf{R}^n	156
7.15.2	Analogous statements for open balls	156
8	Some topics related to differentiability	157
8.1	An integral triangle inequality	157
8.1.1	Vector-valued functions on $[a, b]$	157
8.1.2	Riemann–Stieltjes integrals of vector-valued functions	158
8.1.3	Norms of integrals	158
8.2	A basic Lipschitz estimate	159
8.2.1	Continuously-differentiable vector-valued functions	159
8.2.2	The fundamental theorem of calculus	159
8.2.3	A Lipschitz condition	159
8.2.4	The Lipschitz constant of f	160
8.3	Some partial Lipschitz conditions	161
8.3.1	Functions on cells	161
8.3.2	Using partial Lipschitz conditions	161
8.3.3	Functions on some other subsets of \mathbf{R}^n	162
8.4	Partial derivatives	162
8.4.1	Partial derivatives on cells	163
8.4.2	Partial Lipschitz conditions on cells	163
8.4.3	Continuous differentiability on open subsets of \mathbf{R}^n	163
8.5	A more precise version	163
8.5.1	Real-valued functions on open subsets of \mathbf{R}^n	164
8.6	Directional derivatives	164

8.6.1	A homogeneity property of the directional derivative . . .	165
8.6.2	Some remarks about directional derivatives	165
8.6.3	Linearity in v	165
8.7	Differentiable mappings	166
8.7.1	Some reformulations of differentiability	166
8.7.2	The differential of f at the point x	167
8.7.3	Directional derivatives of differentiable mappings	167
8.7.4	Differentiability of sums and products	167
8.8	Pointwise Lipschitz conditions	168
8.8.1	Defining differentiability using other norms	168
8.8.2	Some pointwise Lipschitz conditions	169
8.8.3	Estimating $\ f'(x)\ _{op}$	169
8.9	The chain rule	170
8.9.1	The precise statement of the chain rule	170
8.9.2	A way to get the chain rule	171
8.9.3	Partial derivatives of compositions	171
8.9.4	Constant functions and linear mappings	171
8.9.5	Squares of differentiable functions	171
8.9.6	Products of differentiable functions	172
8.9.7	Reciprocals of differentiable functions	172
8.9.8	Polynomials and rational functions	172
8.10	Continuous differentiability	172
8.10.1	The case where $\partial_k f(x) = 0$, $1 \leq k \leq n$	173
8.10.2	Reducing to the previous case	173
8.10.3	Another formulation of continuous differentiability	174
8.11	Another basic Lipschitz estimate	174
8.11.1	Continuity of $\ f'(x)\ _{op}$	174
8.11.2	Convex open sets U	175
8.11.3	Estimating $N(f(x) - f(w))$	175
8.11.4	The $m = 1$ case	176
8.12	Some remarks about connectedness	176
8.12.1	Subsets of $X_0 \subseteq X$	176
8.12.2	Disjoint open and closed sets	177
8.12.3	Connected open and closed sets	177
8.13	Locally constant mappings	177
8.13.1	Locally constant mappings and connectedness	178
8.13.2	The condition $f'(x) \equiv 0$	178
8.14	Some local Lipschitz conditions	179
8.14.1	Local Lipschitz conditions and compactness	180
8.14.2	Some Lipschitz conditions on convex subsets of \mathbf{R}^n	180
8.14.3	Some more Lipschitz conditions on subsets of U	181

9	More on differentiable mappings	182
9.1	Continuously-differentiable mappings	182
9.1.1	Using the operator norm	182
9.1.2	Convex open subsets of U	183
9.1.3	A lower bound for $d_N(f(u), f(v))$	183
9.1.4	A lower bound for $N(f'(w)(v))$	183
9.2	Invertible linear mappings	184
9.2.1	The operator norm of A^{-1}	184
9.2.2	Small perturbations of A	185
9.2.3	Estimating $\ B^{-1}\ _{op}$	185
9.3	The inverse function theorem	186
9.3.1	The first part of the inverse function theorem	186
9.3.2	The second part of the inverse function theorem	187
9.4	Proving the inverse function theorem	187
9.4.1	Reducing to the case where $f'(w) = I$ on \mathbf{R}^n	188
9.4.2	The case where $f'(w) = I$	188
9.4.3	A Lipschitz condition for $f(x) - x$ on $B_N(w, r)$	188
9.4.4	Showing the second part	189
9.5	Some remarks about \mathbf{R}^{n+m}	189
9.5.1	Identifying \mathbf{R}^{n+m} with $\mathbf{R}^n \times \mathbf{R}^m$	189
9.5.2	Some related linear mappings	190
9.5.3	The kernel of A	190
9.5.4	Surjectivity of A_1 and A	190
9.5.5	An associated linear mapping \hat{A}	191
9.5.6	Invertibility of \hat{A} and A_1	191
9.5.7	When A is surjective	191
9.6	The implicit function theorem	191
9.6.1	The statement of the implicit function theorem	192
9.6.2	Proving the implicit function theorem	192
9.6.3	Using the inverse function theorem	193
9.6.4	The rest of the argument	193
9.7	Some remarks about determinants	193
9.7.1	Determinants of linear mappings	193
9.7.2	Invertibility and nonzero determinant	194
9.7.3	Continuity of $A \mapsto A^{-1}$	194
9.8	Local embeddings	194
9.8.1	Injectivity of A_1 and A	194
9.8.2	Mappings from open sets in \mathbf{R}^n into \mathbf{R}^{n+m}	195
9.8.3	Invertibility of $f'_1(w)$	195
9.8.4	Getting an invertibility condition	196
9.9	Ranks of linear mappings	196
9.9.1	Some linear subspaces of \mathbf{R}^n	197
9.9.2	Another characterization of the rank	197
9.9.3	Ranks of nearby linear mappings	198
9.9.4	Towards the rank theorem	198
9.10	Some remarks about linear subspaces	198

9.10.1	Sums of linear subspaces	198
9.10.2	Intersections of linear subspaces	199
9.10.3	Kernels of linear mappings	199
9.10.4	The kernel and the rank	199
9.11	Complementary linear subspaces	200
9.11.1	Getting a complementary linear subspace	200
9.11.2	Projections	200
9.11.3	More on projections	201
9.12	The rank theorem	201
9.12.1	The case where $r = \min(m, n)$	202
9.12.2	More on $r = m$ or n	202
9.13	Proving this part	203
9.14	Some more properties of $F \circ H$	204
10	Product spaces and related matters	207
10.1	Products of metric spaces	207
10.1.1	Some metrics on the product	207
10.1.2	Some comparisons between these metrics	208
10.2	Open and closed sets	209
10.2.1	Open and closed balls in the product	209
10.2.2	Products of open sets	210
10.2.3	Closures of products	210
10.2.4	Products of closed sets	210
10.3	Sequences and bounded sets	211
10.3.1	Convergent sequences in the product	211
10.3.2	Cauchy sequences in the product	211
10.3.3	Bounded subsets of the product	211
10.3.4	Total boundedness in the product	212
10.4	Products of compact sets	213
10.4.1	Compactness of products	213
10.4.2	Sequential compactness of products	213
10.4.3	Using completeness and total boundedness	214
10.4.4	Necessity of the compactness of the K_j 's	214
10.5	Some mappings on products	215
10.5.1	Coordinate projections are open mappings	215
10.5.2	Some more partial Lipschitz conditions	215
10.5.3	Functions on product sets	216
10.5.4	Joint and separate continuity	216
10.6	Uniform continuity	216
10.6.1	Uniform continuity along compact sets	217
10.6.2	Uniform continuity and product spaces	217
10.6.3	Uniform continuity and equicontinuity	217
10.6.4	Some compact subsets of a product space	218
10.7	Continuity and integration	218
10.7.1	Integration in x_2	218
10.7.2	Continuity of F	219

10.7.3	Another argument using sequences	219
10.7.4	Uniform continuity of F	219
10.8	Iterated integrals	220
10.8.1	The product of the first l intervals	220
10.8.2	Integrating continuous functions one variable at a time	220
10.8.3	More on iterated integrals	221
10.9	Partitions of intervals	221
10.9.1	Piecewise-linear functions on $[a, b]$	221
10.9.2	Some piecewise-linear approximations	221
10.9.3	Approximating the integral of f	222
10.9.4	Using uniform continuity of f	222
10.10	Partitions and product spaces	223
10.10.1	A partition of $[a_2, b_2]$	223
10.10.2	Integrals in x_2	223
10.10.3	A uniform equicontinuity condition	224
10.11	Partitions and integrals	225
10.11.1	Integrals in x_1	225
10.11.2	Approximating the integral of f in x_1	225
10.11.3	Some iterated integrals	226
10.11.4	Iterated integrals of f	226
10.12	A simpler approximation	227
10.12.1	Approximating some iterated integrals	227
10.12.2	Differences of some iterated integrals	227
10.12.3	Equality of some iterated integrals	228
10.13	Another approximation	228
10.13.1	Iterated integrals as double sums	228
10.13.2	Approximating the summands	229
10.13.3	Equality of the iterated integrals	229
10.14	Partitions of unity	229
10.14.1	Some piecewise-linear functions	230
10.14.2	Getting some partitions of unity	230
10.14.3	Another way to get a partition of unity	231
10.15	Another approximation argument	232
10.15.1	Using a partition of unity	232
10.15.2	An approximation of f	232
10.15.3	Some simple estimates	233
10.15.4	Using these estimates	233
10.16	Some continuous functions	234
10.16.1	Pairs of disjoint closed sets	234
10.16.2	Some uniformly continuous functions	234
10.16.3	Some functions with small support	235
10.16.4	Some functions equal to 1 near a_0	235
10.17	Products and reciprocals	236
10.18	Graphs of mappings	237
10.18.1	Closed graphs in X	238
10.18.2	Compact graphs in X	238

10.19	Semicontinuity	239
10.19.1	Upper and lower semicontinuity	240
10.20	Homeomorphisms between metric spaces	241
10.20.1	A criterion using compactness	241
10.20.2	An argument using closed sets	242
10.21	Graphs and homeomorphisms	242
10.21.1	The case where $X_2 = \mathbf{R}^m$	243
11	Summable functions	245
11.1	Extended real numbers	245
11.1.1	Addition and multiplication	245
11.1.2	Upper and lower bounds	246
11.1.3	Sequences tending to $+\infty$	246
11.1.4	Monotonically increasing sequences	246
11.2	Nonnegative sums	247
11.2.1	Summing nonnegative real-valued functions	247
11.2.2	The case where $X = \mathbf{Z}_+$	247
11.2.3	Equivalence of the two definitions in this case	247
11.2.4	A sometimes convenient extension	248
11.3	Compositions and subsets	248
11.3.1	Sums over subsets	248
11.3.2	Listing terms with a sequence	249
11.3.3	Monotonicity of the sum	249
11.4	Nonnegative summable functions	250
11.4.1	Summability and vanishing at infinity	250
11.4.2	Another property of summable functions	250
11.4.3	More on nonnegative summable functions	251
11.5	Some linearity properties	251
11.5.1	An additivity property	251
11.5.2	The other inequality	252
11.5.3	A related additivity property	253
11.6	ℓ^1 Spaces	253
11.6.1	The ℓ^1 norm	253
11.6.2	The ℓ^1 metric	254
11.6.3	Density of $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ in $\ell^1(X, \mathbf{R})$, $\ell^1(X, \mathbf{C})$	254
11.7	Real-valued summable functions	255
11.7.1	Summing real-valued summable functions	255
11.7.2	A property of the sum	255
11.7.3	Homogeneity of the sum	256
11.7.4	Additivity of the sum	256
11.8	Complex-valued summable functions	256
11.8.1	Summing complex-valued summable functions	257
11.8.2	Additivity for complex-valued functions	257
11.8.3	Homogeneity using complex numbers	257
11.8.4	The absolute value of the sum in the complex case	257
11.9	Some properties of the sum	258

11.9.1	Continuity of the sum on $\ell^1(X, \mathbf{R}), \ell^1(X, \mathbf{C})$	258
11.9.2	More on sums over subsets	259
11.9.3	Another additivity property	260
11.10	Generalized convergence	260
11.10.1	Generalized convergence and summable functions	260
11.10.2	Getting summability	261
11.11	Compositions and sums	262
11.11.1	Sums over X and Y	263
11.11.2	Countably-infinite sets X	263
11.11.3	Relation with infinite series	263
11.12	Completeness of ℓ^1 spaces	264
11.12.1	Cauchy sequences in $\ell^1(X, \mathbf{R}), \ell^1(X, \mathbf{C})$	264
11.12.2	Pointwise convergence	265
11.12.3	Convergence in $\ell^1(X, \mathbf{R}), \ell^1(X, \mathbf{C})$	265
11.13	Monotone convergence	265
11.13.1	Convergence of the sums	266
11.13.2	The proof of convergence	266
11.14	Dominated convergence	267
11.14.1	Getting some sums to be uniformly small	267
11.14.2	Getting some sums to be small when j is large	268
11.14.3	A simple corollary	268
11.15	Nonnegative sums of sums	268
11.15.1	Sums over the E_j 's	269
11.15.2	Estimating the sum over E	269
11.15.3	Estimating the iterated sum	270
11.16	Real and complex sums	270
11.16.1	Some summability conditions	270
11.16.2	Iterated sums	271
11.16.3	Sums over Cartesian products	271
11.17	Square-summable functions	272
11.17.1	The Cauchy–Schwarz inequality in this setting	272
11.17.2	The triangle inequality for $\ \cdot\ _2$	273
11.18	ℓ^2 Spaces	274
11.18.1	Comparison with the supremum norm	274
11.18.2	Comparison with the ℓ^1 norm	274
11.18.3	Density of $c_{00}(X, \mathbf{R}), c_{00}(X, \mathbf{C})$ in $\ell^2(X, \mathbf{R}), \ell^2(X, \mathbf{C})$	275
11.19	Completeness of ℓ^2 spaces	275
11.19.1	Cauchy sequences in $\ell^2(X, \mathbf{R}), \ell^2(X, \mathbf{C})$	276
11.19.2	Convergence in $\ell^2(X, \mathbf{R}), \ell^2(X, \mathbf{C})$	276
11.20	Inner products on ℓ^2 spaces	277
11.20.1	An inner product on $\ell^2(X, \mathbf{C})$	278

12 Some additional topics	279
12.1 Lebesgue measure and integration	279
12.1.1 The bounded convergence theorem	279
12.1.2 Continuity almost everywhere	280
12.1.3 Monotone functions	280
12.1.4 Lipschitz functions	280
12.2 Banach spaces	281
12.2.1 Norms on vector spaces	281
12.2.2 Some remarks about convergent sequences	282
12.2.3 Linear subspaces	282
12.3 Hilbert spaces	282
12.3.1 The complex case	283
12.3.2 Norms associated to inner products	283
12.4 Infinite series in Banach spaces	284
12.4.1 Some basic properties of infinite series in V	284
12.4.2 The Cauchy criterion	285
12.4.3 Absolute convergence with respect to the norm	285
12.4.4 Orthogonal vectors in an inner product space	286
12.5 Bounded linear mappings	286
12.5.1 Some related characterizations	287
12.5.2 The operator norm of a bounded linear mapping	287
12.5.3 Spaces of bounded linear mappings	288
12.5.4 Compositions of bounded linear mappings	288
12.6 Dual spaces	288
12.6.1 Linear functionals on $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$	288
12.6.2 Bounded linear functionals	289
12.6.3 Using an inner product	289
12.6.4 A part of the Hahn–Banach theorem	290
12.7 Some bounded linear functionals	290
12.7.1 Some more examples	290
Appendix	292
A Some more on mappings, metrics, and norms	292
A.1 A nice inequality	292
A.1.1 Snowflake transforms	292
A.2 Some more Lipschitz conditions	293
A.2.1 Lipschitz functions of order $\alpha > 1$ on \mathbf{R}	294
A.3 Another nice inequality	294
A.3.1 The limit of $\ f\ _p$ as $p \rightarrow \infty$	295
A.3.2 A related metric for $p \leq 1$	295
A.4 Some functions on $[0, 1]$	296
A.4.1 Some related sequences of functions	297
A.5 Some Cauchy sequences	297
A.5.1 Another sequence of functions	298
A.5.2 Some functions on \mathbf{Z}_+	298

A.6	Norms and convexity	299
A.6.1	The definition of a seminorm	299
A.6.2	Some norms on $c_{00}(X, \mathbf{R})$	299
A.6.3	Complex vector spaces	300
A.6.4	More on seminorms	300
A.7	Path-connected sets	300
A.7.1	The definition of path connectedness	301
A.7.2	Connected open sets in \mathbf{R}^n	301
A.7.3	Closures of connected sets	301
A.8	Bounded vector-valued functions	302
A.8.1	Using a norm N on \mathbf{R}^m or \mathbf{C}^m	302
A.8.2	Supremum norms associated to N	303
A.8.3	Continuous vector-valued functions	303
A.9	Some remarks about n -dimensional volume	304
A.9.1	Volumes of cells	304
A.9.2	Translations and dilations	304
A.10	Volumes and Lipschitz mappings	305
A.10.1	A basic estimate for volumes	305
A.11	Compact support	306
A.11.1	Some properties of the support	306
A.11.2	Compactly supported functions	307
A.11.3	Compactly supported continuous functions	307
A.12	Local compactness	308
A.12.1	Local compactness and continuous functions	308
A.13	Another vanishing condition	309
A.13.1	Some equivalent reformulations	309
A.13.2	Continuous functions vanishing at infinity	310
A.13.3	A uniform convergence property	310
A.13.4	Some additional properties	311
	Bibliography	312
	Index	328

Chapter 1

Some mappings, metrics, and norms

1.1 Lipschitz mappings

Let $(X, d_X(\cdot, \cdot))$ and $(Y, d_Y(\cdot, \cdot))$ be metric spaces. A mapping f from X into Y is said to be *Lipschitz* if there is a nonnegative real number C such that

$$(1.1.1) \quad d_Y(f(x), f(w)) \leq C d_X(x, w)$$

for every $x, w \in X$. It is easy to see that Lipschitz mappings are uniformly continuous. Note that (1.1.1) holds with $C = 0$ if and only if f is constant on X .

1.1.1 A simple relationship between metrics

Let X be a set, and let $d(\cdot, \cdot)$ and $d'(\cdot, \cdot)$ be metrics on X . Consider the condition that there be a nonnegative real number C such that

$$(1.1.2) \quad d'(x, w) \leq C d(x, w)$$

for every $x, w \in X$. This is the same as (1.1.1), with $d_X = d$, $Y = X$, $d_Y = d'$, and f taken to be the identity mapping on X .

1.1.2 Square roots of metrics

Let $(X, d(\cdot, \cdot))$ be a metric space, and put

$$(1.1.3) \quad \rho(x, w) = \sqrt{d(x, w)}$$

for every $x, w \in X$. One can check that this defines a metric on X as well. This corresponds to the second part of Exercise 11 at the end of Chapter 2 in [210], when we start with the standard Euclidean metric on the real line.

One can check that the identity mapping on X is uniformly continuous as a mapping from X equipped with $d(\cdot, \cdot)$ into X equipped with $\rho(x, w)$. Similarly, one can check that the identity mapping on X is uniformly continuous as a mapping from X equipped with $\rho(\cdot, \cdot)$ into X equipped with $d(\cdot, \cdot)$.

1.2 Lipschitz conditions on \mathbf{R}

Let E be a nonempty subset of the real line \mathbf{R} , and let f be a real-valued function on E . Note that f is Lipschitz on E with respect to the standard Euclidean metric on \mathbf{R} and its restriction to E if and only if there is a nonnegative real number C such that

$$(1.2.1) \quad |f(x) - f(w)| \leq C|x - w|$$

for every $x, w \in E$. Here $|t|$ denotes the usual absolute value of a real number t .

Of course, (1.2.1) holds automatically when $x = w$. If $x \neq w$, then (1.2.1) is the same as saying that

$$(1.2.2) \quad \frac{|f(x) - f(w)|}{|x - w|} \leq C.$$

1.2.1 The definition of the derivative

Suppose that $x \in E$ is a limit point of E . The *derivative* of f at x is defined as usual by

$$(1.2.3) \quad f'(x) = \lim_{\substack{w \in E \\ w \rightarrow x}} \frac{f(w) - f(x)}{w - x},$$

when the limit on the right exists. In this case, f is said to be *differentiable* at x , as a function on E . If f is differentiable at x , and (1.2.1) holds for all $w \in E$, or at least when $w \in E$ is sufficiently close to x , then one can check that

$$(1.2.4) \quad |f'(x)| \leq C.$$

If f is differentiable at x , then f is continuous at x , as a function on E , by a standard argument. More precisely, if C is a real number such that

$$(1.2.5) \quad |f'(x)| < C,$$

then one can verify that (1.2.1) holds for all $w \in E$ that are sufficiently close to x .

1.2.2 Using the mean value theorem

Let a and b be real numbers with $a < b$, and suppose for the moment that E is the corresponding open interval (a, b) in \mathbf{R} . We may also allow $a = -\infty$ or $b = +\infty$ here, so that E could be the real line, or an open half-line in \mathbf{R} . Suppose for the moment that

$$(1.2.6) \quad f \text{ is differentiable at every point in } E,$$

and that there is a nonnegative real number C such that (1.2.4) holds for every $x \in E$. Under these conditions, the mean value theorem implies that (1.2.1) holds for every $x, w \in E$.

Let a and b be real numbers with $a < b$ again, and suppose now that E is the corresponding closed interval $[a, b]$ in \mathbf{R} . Suppose for the moment again that

(1.2.7) f is continuous on $[a, b]$, and differentiable at every point in (a, b) .

If there is a nonnegative real number C such that (1.2.4) holds for every x in (a, b) , then the mean-value theorem implies that (1.2.1) holds for every $x, w \in E$.

Of course, there are analogous statements when E is a half-open, half-closed interval in \mathbf{R} , or a closed half-line in \mathbf{R} .

See [18, 47, 58] for some related perspectives on the mean value theorem. Some additional results related to the mean value theorem can be found in [138, 244].

Some aspects of calculus on the rationals are discussed in [140, 142]. Some more connections between Lipschitz conditions and derivatives will be considered in Chapter 8.

1.3 Norms on \mathbf{R}^n

Let n be a positive integer. Remember that \mathbf{R}^n is the space of ordered n -tuples $x = (x_1, \dots, x_n)$ such that $x_j \in \mathbf{R}$ for each $j = 1, \dots, n$. Addition can be defined on \mathbf{R}^n coordinatewise, so that

$$(1.3.1) \quad x + y = (x_1 + y_1, \dots, x_n + y_n)$$

for every $x, y \in \mathbf{R}^n$. Similarly, if $t \in \mathbf{R}$ and $x \in \mathbf{R}^n$, then

$$(1.3.2) \quad tx = (tx_1, \dots, tx_n)$$

defines another element of \mathbf{R}^n . Using these definitions of addition and scalar multiplication, \mathbf{R}^n becomes a vector space over the real numbers. Although we shall not discuss the formal definition of a vector space here, the relevant notions will hopefully be clear in the examples. Note that we shall use 0 to refer to the element of \mathbf{R}^n whose coordinates are equal to the real number 0 , which will hopefully also be clear from the context.

1.3.1 The definition of a norm

A nonnegative real-valued function N on \mathbf{R}^n is said to define a *norm* on \mathbf{R}^n if it satisfies the following three conditions. First, $N(x) = 0$ if and only if $x = 0$. Second,

$$(1.3.3) \quad N(tx) = |t| N(x)$$

for every $t \in \mathbf{R}$ and $x \in \mathbf{R}^n$. Third,

$$(1.3.4) \quad N(x + y) \leq N(x) + N(y)$$

for every $x, y \in \mathbf{R}^n$, which is the *triangle inequality* for norms.

1.3.2 Some basic examples of norms

The standard *Euclidean norm* is defined by

$$(1.3.5) \quad \|x\|_2 = \left(\sum_{j=1}^n x_j^2 \right)^{1/2}$$

for every $x \in \mathbf{R}^n$. Of course, this uses the nonnegative square root on the right side of the equation. It is easy to see that this satisfies the first two requirements of a norm mentioned in the preceding paragraph. The triangle inequality is more complicated, and can be obtained from the Cauchy–Schwarz inequality. See Theorem 1.37 on p16 of [210].

One can check directly that

$$(1.3.6) \quad \|x\|_1 = \sum_{j=1}^n |x_j|$$

defines a norm on \mathbf{R}^n . (Exercise.) Similarly,

$$(1.3.7) \quad \|x\|_\infty = \max(|x_1|, \dots, |x_n|)$$

defines a norm on \mathbf{R}^n , where more precisely the right side is the maximum of $|x_1|, \dots, |x_n|$. In particular, if $x, y \in \mathbf{R}^n$, then

$$(1.3.8) \quad \begin{aligned} \|x + y\|_\infty &= \max(|x_1 + y_1|, \dots, |x_n + y_n|) \\ &\leq \max(|x_1| + |y_1|, \dots, |x_n| + |y_n|) \leq \|x\|_\infty + \|y\|_\infty, \end{aligned}$$

using the triangle inequality for the absolute value function on \mathbf{R} in the second step.

1.3.3 Metrics associated to norms

If N is any norm on \mathbf{R}^n , then

$$(1.3.9) \quad d_N(x, y) = N(x - y)$$

defines a metric on \mathbf{R}^n . Indeed, the first requirement of a norm ensures that (1.3.9) is equal to 0 if and only if $x = y$. The homogeneity condition (1.3.3) implies that (1.3.9) is symmetric in x and y , by taking $t = -1$ in (1.3.3). The triangle inequality for (1.3.9) as a metric on \mathbf{R}^n follows from the triangle inequality (1.3.4) for N as a norm on \mathbf{R}^n .

The metric

$$(1.3.10) \quad d_2(x, y) = \|x - y\|_2$$

associated to the standard Euclidean norm (1.3.5) is the standard *Euclidean metric* on \mathbf{R}^n . Let

$$(1.3.11) \quad d_1(x, y) = \|x - y\|_1$$

and

$$(1.3.12) \quad d_\infty(x, y) = \|x - y\|_\infty$$

be the metrics on \mathbf{R}^n corresponding to the norms (1.3.6) and (1.3.7), respectively. If $n = 1$, then the norms (1.3.5), (1.3.6), and (1.3.7) reduce to the absolute value function on \mathbf{R} , and the corresponding metrics are the same as the standard Euclidean metric on \mathbf{R} .

1.3.4 The Cauchy–Schwarz inequality for finite sums

If $a_1, \dots, a_n, b_1, \dots, b_n$ are nonnegative real numbers, then

$$(1.3.13) \quad \sum_{j=1}^n a_j b_j \leq \left(\sum_{j=1}^n a_j^2 \right)^{1/2} \left(\sum_{j=1}^n b_j^2 \right)^{1/2}.$$

This is a version of the *Cauchy–Schwarz inequality*. This is often formulated a bit differently for arbitrary real or complex numbers, as in Theorem 1.35 on p15 of [210]. This formulation is included in the other one, by restricting one's attention to nonnegative real numbers. The other formulation can also be obtained from this one, by applying (1.3.13) to the absolute values of the given real or complex numbers.

1.4 Norms on \mathbf{C}^n

Let \mathbf{C} be the complex plane, as usual, and let n be a positive integer again. As before, \mathbf{C}^n is the set of ordered n -tuples of complex numbers. Addition can be defined on \mathbf{C}^n coordinatewise, as in (1.3.1). One can also multiply an element of \mathbf{C}^n by a complex number coordinatewise, as in (1.3.2). In this way, \mathbf{C}^n becomes a vector space over the complex numbers.

A nonnegative real-valued function N on \mathbf{C}^n is said to be a *norm* on \mathbf{C}^n if it satisfies the same type of conditions as in the previous section. More precisely, the first and third conditions are the same as before. In this situation, the homogeneity condition (1.3.3) should hold for all complex numbers t and elements of \mathbf{C}^n , where $|t|$ is the usual absolute value function on \mathbf{C} .

The standard *Euclidean norm* is defined on \mathbf{C}^n by

$$(1.4.1) \quad \|z\|_2 = \left(\sum_{j=1}^n |z_j|^2 \right)^{1/2},$$

where $|z_j|$ is the absolute value of $z_j \in \mathbf{C}$ for each $j = 1, \dots, n$. The triangle inequality for (1.4.1) can be reduced to the real case, by taking the absolute values of the coordinates of elements of \mathbf{C}^n to get n -tuples of nonnegative real numbers. This argument uses the triangle inequality for the absolute value function on \mathbf{C} , which is the same as the $n = 1$ case.

The triangle inequality for (1.4.1) on \mathbf{C}^n can also be obtained from the Cauchy–Schwarz inequality, using an argument analogous to the one in the real

case. As before, it is easy to verify the other two requirements for (1.4.1) to be a norm on \mathbf{C}^n directly from the definition.

Alternatively, the complex plane can be identified with \mathbf{R}^2 , using the real and imaginary parts of a complex number. Using this identification, the absolute value of a complex number corresponds to the standard Euclidean norm on \mathbf{R}^2 .

Similarly, \mathbf{C}^n can be identified with \mathbf{R}^{2n} , using the real and imaginary parts of the n coordinates of an element of \mathbf{C}^n . Using this identification, the standard Euclidean norm (1.4.1) on \mathbf{C}^n corresponds exactly to the standard Euclidean norm on \mathbf{R}^{2n} . This permits one to get the triangle inequality for (1.4.1) on \mathbf{C}^n from the triangle inequality for the standard Euclidean norm on \mathbf{R}^{2n} , because addition on \mathbf{C}^n corresponds exactly to addition on \mathbf{R}^{2n} with respect to this identification.

As before, one can verify directly that

$$(1.4.2) \quad \|z\|_1 = \sum_{j=1}^n |z_j|$$

and

$$(1.4.3) \quad \|z\|_\infty = \max(|z_1|, \dots, |z_n|)$$

define norms on \mathbf{C}^n as well. Of course, \mathbf{R}^n may be considered as a subset of \mathbf{C}^n , because \mathbf{R} is contained in \mathbf{C} . The restrictions of (1.4.2) and (1.4.3) to $z \in \mathbf{R}^n$ are the same as the corresponding norms defined on \mathbf{R}^n in the previous section. Similarly, the restriction of (1.4.1) to $z \in \mathbf{R}^n$ is the same as the standard Euclidean norm on \mathbf{R}^n .

If N is any norm on \mathbf{C}^n , then

$$(1.4.4) \quad d_N(z, w) = N(z - w)$$

defines a metric on \mathbf{C}^n , for the same reasons as in the real case. The standard *Euclidean metric* on \mathbf{C}^n is the metric

$$(1.4.5) \quad d_2(z, w) = \|z - w\|_2$$

associated to the standard Euclidean norm (1.4.1). Similarly, let

$$(1.4.6) \quad d_1(z, w) = \|z - w\|_1$$

and

$$(1.4.7) \quad d_\infty(z, w) = \|z - w\|_\infty$$

be the metrics on \mathbf{C}^n associated to the norms (1.4.2) and (1.4.3), respectively. If $n = 1$, then the norms (1.4.1), (1.4.2), and (1.4.3) reduce to the absolute value function on \mathbf{C} , so that the corresponding metrics are the same as the standard Euclidean metric on \mathbf{C} .

The restriction of any norm N on \mathbf{C}^n to \mathbf{R}^n defines a norm on \mathbf{R}^n . In this case, the restriction of (1.4.4) to $z, w \in \mathbf{R}^n$ is the same as the metric on \mathbf{R}^n associated to the restriction of N to \mathbf{R}^n . In particular, the restrictions of (1.4.5), (1.4.6), and (1.4.7) to $z, w \in \mathbf{R}^n$ are the same as the corresponding metrics defined on \mathbf{R}^n in the previous section.

1.4.1 A larger family of norms

Let p be a positive real number, and put

$$(1.4.8) \quad \|z\|_p = \left(\sum_{j=1}^n |z_j|^p \right)^{1/p}$$

for every $z \in \mathbf{C}^n$. It is easy to see that this satisfies the positivity and homogeneity requirements of a norm. If $p \geq 1$, then it is well known that (1.4.8) satisfies the triangle inequality, and hence defines a norm on \mathbf{C}^n . This is a version of *Minkowski's inequality* for sums. One approach to this is discussed in Section A.6.

Of course, (1.4.8) is the same as (1.4.1) when $p = 2$, and it is the same as (1.4.2) when $p = 1$. If $n = 1$, then (1.4.8) reduces to the absolute value function on \mathbf{C} . If $0 < p < 1$ and $n \geq 2$, then one can check that (1.4.8) does not satisfy the triangle inequality. This is related to the lack of convexity of the corresponding balls in \mathbf{C}^n , as in Section A.6. There are analogous statements for the restriction of (1.4.8) to \mathbf{R}^n .

It is not too difficult to show that (1.4.8) tends to $\|z\|_\infty$ as $p \rightarrow \infty$ for every $z \in \mathbf{C}^n$, as in Section A.3.

1.5 Some basic inequalities

Let n be a positive integer, and let $z \in \mathbf{C}^n$ be given. It is easy to see that

$$(1.5.1) \quad \|z\|_\infty \leq \|z\|_2, \quad \|z\|_2 \leq \|z\|_1,$$

directly from the definitions of these norms in the previous section.

Similarly,

$$(1.5.2) \quad \|z\|_2^2 = \sum_{j=1}^n |z_j|^2 \leq \|z\|_\infty \sum_{j=1}^n |z_j| = \|z\|_\infty \|z\|_1 \leq \|z\|_1^2,$$

so that

$$(1.5.3) \quad \|z\|_2 \leq \|z\|_1.$$

It follows that the corresponding metrics satisfy

$$(1.5.4) \quad d_\infty(z, w) \leq d_2(z, w) \leq d_1(z, w)$$

for every $z, w \in \mathbf{C}^n$. There are some analogous statements for (1.4.8) for all $p > 0$, which are discussed in Section A.3.

1.5.1 Inequalities in the other direction

In the other direction, we have that

$$(1.5.5) \quad \|z\|_2 \leq n^{1/2} \|z\|_\infty$$

and

$$(1.5.6) \quad \|z\|_1 \leq n \|z\|_\infty$$

for every $z \in \mathbf{C}^n$. One can also check that

$$(1.5.7) \quad \|z\|_1 \leq n^{1/2} \|z\|_2$$

for every $z \in \mathbf{C}^n$, using the Cauchy-Schwarz inequality (1.3.13). This implies that

$$(1.5.8) \quad d_2(z, w) \leq n^{1/2} d_\infty(z, w),$$

$$(1.5.9) \quad d_1(z, w) \leq n d_\infty(z, w),$$

and

$$(1.5.10) \quad d_1(z, w) \leq n^{1/2} d_2(z, w)$$

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

1.5.2 Some consequences of these inequalities

Using these simple relationships, we get that $d_1(z, w)$, $d_2(z, w)$, and $d_\infty(z, w)$ have many of the same properties on \mathbf{C}^n . They determine the same collections of open sets, closed sets, compact sets, and bounded sets, for instance. They also determine the same limit points of subsets of \mathbf{C}^n , convergent sequences in \mathbf{C}^n , and Cauchy sequences.

Using (1.5.4), it is easy to see that the identity mapping on \mathbf{C}^n is Lipschitz as a mapping from \mathbf{C}^n equipped with $d_1(z, w)$ into \mathbf{C}^n equipped with $d_2(z, w)$, and from \mathbf{C}^n equipped with $d_2(z, w)$ into \mathbf{C}^n equipped with $d_\infty(z, w)$. Similarly, the identity mapping on \mathbf{C}^n is Lipschitz as a mapping from \mathbf{C}^n equipped with $d_1(z, w)$ into \mathbf{C}^n equipped with $d_\infty(z, w)$. We can use (1.5.8) to get that the identity mapping on \mathbf{C}^n is Lipschitz as a mapping from \mathbf{C}^n equipped with $d_\infty(z, w)$ into \mathbf{C}^n equipped with $d_2(z, w)$, and (1.5.10) implies that the identity mapping on \mathbf{C}^n is Lipschitz as a mapping from \mathbf{C}^n equipped with $d_2(z, w)$ into \mathbf{C}^n equipped with $d_1(z, w)$. One can use (1.5.9) to get that the identity mapping on \mathbf{C}^n is Lipschitz as a mapping from \mathbf{C}^n equipped with $d_\infty(z, w)$ into \mathbf{C}^n equipped with $d_1(z, w)$.

Of course, there are analogous statements for the restrictions of these metrics to \mathbf{R}^n .

1.5.3 Estimates for arbitrary norms

If N is any norm on \mathbf{R}^n or \mathbf{C}^n , then one can show that N is bounded by a constant times the standard Euclidean norm, or equivalently by a constant times either of the norms $\|\cdot\|_1$, $\|\cdot\|_\infty$. To see this, one can express any element of \mathbf{R}^n or \mathbf{C}^n as a linear combination of the standard basis vectors, to estimate N in terms of the absolute values of the coordinates of the given vector. This is discussed further in Section 6.10.

One can use this to show that N is continuous as a real-valued function on \mathbf{R}^n or \mathbf{C}^n , as appropriate, with respect to the standard Euclidean metric. It follows that N attains its minimum on the unit sphere in \mathbf{R}^n or \mathbf{C}^n , as appropriate, with respect to the standard Euclidean metric again, by the extreme value theorem. Because the infimum is positive, by definition of a norm, one can verify that the standard Euclidean norm is bounded by a constant times N on \mathbf{R}^n or \mathbf{C}^n , as appropriate. This is discussed in more detail in Section 6.11.

1.6 Functions with finite support

Let X be a nonempty set, and let f be a real or complex-valued function on X . The *support* of f is defined to be the subset of X given by

$$(1.6.1) \quad \text{supp } f = \{x \in X : f(x) \neq 0\}.$$

Let $c_{00}(X, \mathbf{R})$ be the space of real-valued functions on X whose support has only finitely many elements, and let $c_{00}(X, \mathbf{C})$ be the space of complex-valued functions on X with finite support.

If f and g are real or complex-valued functions on X with finite support, then their sum $f(x) + g(x)$ also defines a real or complex-valued function on X with finite support. Similarly, if f is a real or complex-valued function on X with finite support, and t is a real or complex number, as appropriate, then $tf(x)$ has finite support in X . More precisely, $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$ are vector spaces over the real and complex numbers, respectively, with respect to pointwise addition and scalar multiplication of functions. These may be considered as linear subspaces of the spaces of all real or complex-valued functions on X , respectively.

Of course, if X has only finitely many elements, then every real or complex-valued function on X automatically has finite support. Let n be a positive integer, and suppose for the moment that

$$(1.6.2) \quad X = \{1, \dots, n\}$$

is the set of positive integers from 1 to n . In this case, $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$ can be identified with \mathbf{R}^n and \mathbf{C}^n , respectively.

Similarly, if X is the set \mathbf{Z}_+ of positive integers, then a real or complex-valued function on X corresponds to an infinite sequence of real or complex numbers. Thus $c_{00}(\mathbf{Z}_+, \mathbf{R})$ and $c_{00}(\mathbf{Z}_+, \mathbf{C})$ can be identified with the spaces of infinite sequences of real or complex numbers for which all but finitely many terms are equal to 0, respectively.

1.6.1 Norms on $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$

As usual, a nonnegative real-valued function N on $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$ is said to be a *norm* if it satisfies the following three conditions. First, $N(f) = 0$

if and only if $f = 0$. Second, if $f \in c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$ and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then

$$(1.6.3) \quad N(tf) = |t|N(f).$$

Third,

$$(1.6.4) \quad N(f+g) \leq N(f) + N(g)$$

for every $f, g \in c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$, as appropriate. In this case,

$$(1.6.5) \quad d_N(f, g) = N(f - g)$$

defines a metric on $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$, as appropriate.

Let f be a real or complex-valued function on X . If A is a nonempty finite subset of X , then

$$(1.6.6) \quad \sum_{x \in A} f(x)$$

can be defined as a real or complex number, as appropriate. Suppose that f has finite support in X , and observe that the finite sums (1.6.6) are all the same when $\text{supp } f \subseteq A$. This permits us to define the sum

$$(1.6.7) \quad \sum_{x \in X} f(x)$$

as a real or complex number, as appropriate, as the value of (1.6.6) when A is a nonempty finite subset of X that contains the support of f .

If f is a real or complex-valued function on X with finite support, then

$$(1.6.8) \quad \|f\|_1 = \sum_{x \in X} |f(x)|$$

is defined as a nonnegative real number, as in the previous paragraph. Similarly,

$$(1.6.9) \quad \|f\|_2 = \left(\sum_{x \in X} |f(x)|^2 \right)^{1/2}$$

is defined as a nonnegative real number, where the sum on the right is defined as before. We can also put

$$(1.6.10) \quad \|f\|_\infty = \max_{x \in X} |f(x)|,$$

where the maximum of $|f(x)|$ over $x \in X$ is clearly attained in this situation. One can check that these define norms on $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$. In particular, the triangle inequality for (1.6.9) reduces to the analogous statement for \mathbf{R}^n and \mathbf{C}^n , mentioned in Subsection 1.3.2 and Section 1.4.

Using these norms, we get metrics

$$(1.6.11) \quad d_1(f, g) = \|f - g\|_1,$$

$$(1.6.12) \quad d_2(f, g) = \|f - g\|_2,$$

and

$$(1.6.13) \quad d_\infty(f, g) = \|f - g\|_\infty$$

on $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$. If X is as in (1.6.2) for some positive integer n , then the norms mentioned in the previous paragraph correspond to the analogous norms defined on \mathbf{R}^n and \mathbf{C}^n in Subsection 1.3.2 and Section 1.4. Similarly, these metrics correspond to the analogous metrics defined earlier on \mathbf{R}^n and \mathbf{C}^n in this case.

1.6.2 Some inequalities between these norms

If f is a real or complex-valued function on X with finite support, then

$$(1.6.14) \quad \|f\|_\infty \leq \|f\|_2 \leq \|f\|_1.$$

This follows from (1.5.1) and (1.5.3), since it is enough to look at the finitely many elements of X in the support of f . This implies that

$$(1.6.15) \quad d_\infty(f, g) \leq d_2(f, g) \leq d_1(f, g)$$

for all real and complex-valued functions f and g on X with finite support.

It follows that convergent sequences in $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$ with respect to d_1 are also convergent with respect to d_2 and d_∞ , with the same limit. Similarly, convergent sequences with respect to d_2 are convergent with respect to d_∞ , with the same limit.

We also have that open sets in $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$ with respect to d_∞ are open sets with respect to d_1 and d_2 . Similarly, open sets with respect to d_2 are open sets with respect to d_1 .

There are analogous statements for other properties, which may be considered in terms of the identity mappings on $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$, using these metrics on the domain and range. More precisely, the identity mappings on $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$ are Lipschitz with respect to $d_2(f, g)$ on the domain and $d_\infty(f, g)$ on the range, because of the first inequality in (1.6.15). Similarly, the identity mappings on $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$ are Lipschitz with respect to $d_1(f, g)$ on the domain and $d_2(f, g)$ on the range, because of the second inequality in (1.6.15). The identity mappings on $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$ are also Lipschitz with respect to $d_1(f, g)$ on the domain and $d_\infty(f, g)$ on the range.

1.6.3 Cauchy sequences in $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$

The definition of a Cauchy sequence in a metric space is reviewed in the next section. Let $\{f_j\}_{j=1}^\infty$ be a sequence of real or complex-valued functions on X with finite support. If $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence with respect to d_1 , then $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence with respect to d_2 and d_∞ , because of (1.6.15). Similarly, if $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence with respect to d_2 , then $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence with respect to d_∞ . Some classes of examples of Cauchy sequences like these are discussed in Subsection A.5.2.

If X has infinitely many elements, then it is not too difficult to see that neither $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$ is complete with respect to any of the metrics d_1 , d_2 , or d_∞ , using examples of Cauchy sequences like those in Subsection A.5.2. To get complete metric spaces, one can consider additional real and complex-valued functions on X , depending on the metric d_1 , d_2 , or d_∞ . This is related to Chapter 2 for d_1 , d_2 when $X = \mathbf{Z}_+$, and to Chapter 11 for arbitrary X . This is also related to Sections 1.13 and 2.5 for d_∞ .

1.7 Cauchy sequences and completeness

Let $(X, d_X(\cdot, \cdot))$ be a metric space. Remember that a sequence $\{x_j\}_{j=1}^\infty$ of elements of X is said to be a *Cauchy sequence* with respect to $d_X(\cdot, \cdot)$ if for every $\epsilon > 0$ there is a positive integer L such that

$$(1.7.1) \quad d_X(x_j, x_l) < \epsilon$$

for every $j, l \geq L$. It is not difficult to verify that

$$(1.7.2) \quad \text{convergent sequences in } X \text{ are Cauchy sequences.}$$

If

$$(1.7.3) \quad \text{every Cauchy sequence in } X \text{ converges to an element of } X,$$

then X is said to be *complete* with respect to $d_X(\cdot, \cdot)$. It is well known that \mathbf{R} and \mathbf{C} are complete with respect to their standard Euclidean metrics.

1.7.1 Cauchy sequences in subsets of metric spaces

Let E be a subset of X , and remember that the restriction of $d_X(x, w)$ to $x, w \in E$ defines a metric on E . If $\{x_j\}_{j=1}^\infty$ is a sequence of elements of E , then it is easy to see that

$$(1.7.4) \quad \{x_j\}_{j=1}^\infty \text{ is a Cauchy sequence as a sequence of elements of } E$$

if and only if

$$(1.7.5) \quad \{x_j\}_{j=1}^\infty \text{ is a Cauchy sequence as a sequence of elements of } X,$$

with respect to $d_X(\cdot, \cdot)$ and its restriction to E .

If X is complete with respect to $d_X(\cdot, \cdot)$, and if E is a closed set in X , then

$$(1.7.6) \quad E \text{ is complete with respect to the restriction of } d_X(\cdot, \cdot) \text{ to } E.$$

More precisely, if $\{x_j\}_{j=1}^\infty$ is a Cauchy sequence of elements of E , then $\{x_j\}_{j=1}^\infty$ is a Cauchy sequence in X . This implies that $\{x_j\}_{j=1}^\infty$ converges to some $x \in X$, because X is complete. We also have that $x \in E$, because E is a closed set in X . It follows that $\{x_j\}_{j=1}^\infty$ converges to x in E , with respect to the restriction of $d_X(\cdot, \cdot)$ to E .

1.7.2 Completeness of subsets

Suppose now that E is complete with respect to the restriction of $d_X(\cdot, \cdot)$ to E , and let us check that

$$(1.7.7) \quad E \text{ is a closed set in } X.$$

Let $x \in X$ be a limit point of E , which implies that there is a sequence $\{x_j\}_{j=1}^{\infty}$ of elements of E that converges to x in X . Note that $\{x_j\}_{j=1}^{\infty}$ is a Cauchy sequence as a sequence of elements of X , and hence as a sequence of elements of E . Because E is complete, there is an $x' \in E$ such that $\{x_j\}_{j=1}^{\infty}$ converges to x' with respect to the restriction of $d_X(\cdot, \cdot)$ to E . Of course, $\{x_j\}_{j=1}^{\infty}$ converges to x' in X as well, so that $x = x'$, and thus $x \in E$.

1.7.3 Cauchy sequences and uniform continuity

Let $(Y, d_Y(\cdot, \cdot))$ be another metric space, and let f be a uniformly continuous mapping from X into Y . If $\{x_j\}_{j=1}^{\infty}$ is a Cauchy sequence of elements of X , then one can check that

$$(1.7.8) \quad \{f(x_j)\}_{j=1}^{\infty} \text{ is a Cauchy sequence in } Y.$$

This is the first part of Exercise 11 at the end of Chapter 4 in [210]. If Y is complete, then it follows that $\{f(x_j)\}_{j=1}^{\infty}$ converges in Y . If $\{x_j\}_{j=1}^{\infty}$ converges to an element x of X , then continuity of f at x implies that $\{f(x_j)\}_{j=1}^{\infty}$ converges to $f(x)$ in Y .

1.8 Pointwise and uniform convergence

Let X be a set, and let $(Y, d_Y(\cdot, \cdot))$ be a metric space. Also let $\{f_j\}_{j=1}^{\infty}$ be a sequence of mappings from X into Y , and let f be another mapping from X into Y . We say that $\{f_j\}_{j=1}^{\infty}$ converges to f *pointwise* on X if for every $x \in X$, $\{f_j(x)\}_{j=1}^{\infty}$ converges to $f(x)$ in Y . This means that for every $x \in X$ and $\epsilon > 0$ there is a positive integer L such that

$$(1.8.1) \quad d_Y(f_j(x), f(x)) < \epsilon$$

for every $j \geq L$.

We say that $\{f_j\}_{j=1}^{\infty}$ converges to f *uniformly* on X if for every $\epsilon > 0$ there is a positive integer L such that (1.8.1) holds for every $x \in X$ and $j \geq L$. Note that uniform convergence implies pointwise convergence. If X has only finitely many elements, and $\{f_j\}_{j=1}^{\infty}$ converges to f pointwise on X , then one can check that $\{f_j\}_{j=1}^{\infty}$ converges to f uniformly on X .

1.8.1 An example on $[0, 1]$

As an example, let us take X to be the closed unit interval $[0, 1]$ in the real line, and $Y = \mathbf{R}$ with the standard metric. Put

$$(1.8.2) \quad f_j(x) = x^j$$

for each positive integer j and $0 \leq x \leq 1$. In this case,

$$(1.8.3) \quad \begin{aligned} \lim_{j \rightarrow \infty} f_j(x) &= 0 \quad \text{when } 0 \leq x < 1 \\ &= 1 \quad \text{when } x = 1. \end{aligned}$$

However, $\{f_j\}_{j=1}^{\infty}$ does not converge uniformly on $[0, 1]$, because for each positive integer j we have that x^j is as close to 1 as we want when x is sufficiently close to 1. If r is a positive real number with $r < 1$, then $\{f_j\}_{j=1}^{\infty}$ does converge to 0 uniformly on $[0, r]$.

1.8.2 Uniform convergence and continuity

Now let (X, d_X) be a metric space, and let (Y, d_Y) be a metric space again too. Also let

$$(1.8.4) \quad \begin{aligned} \{f_j\}_{j=1}^{\infty} &\text{ be a sequence of mappings from } X \text{ into } Y \\ &\text{that converges uniformly to a mapping } f \text{ from } X \text{ into } Y, \end{aligned}$$

and let $x \in X$ be given. If

$$(1.8.5) \quad f_j \text{ is continuous at } x \text{ for every } j \geq 1,$$

then it is well known that

$$(1.8.6) \quad f \text{ is continuous at } x$$

as well. To see this, let $\epsilon > 0$ be given. Because $\{f_j\}_{j=1}^{\infty}$ converges uniformly to f on X , there is an $L \in \mathbf{Z}_+$ such that

$$(1.8.7) \quad d_Y(f_j(w), f(w)) < \epsilon/3$$

for every $j \geq L$ and $w \in X$. In particular, this holds at x , so that

$$(1.8.8) \quad d_Y(f_j(x), f(x)) < \epsilon/3$$

for every $j \geq L$. Because f_L is continuous at x , there is a $\delta_L > 0$ such that

$$(1.8.9) \quad d_Y(f_L(x), f_L(w)) < \epsilon/3$$

for every $w \in X$ with $d_X(x, w) < \delta_L$. Observe that

$$(1.8.10) \quad \begin{aligned} d_Y(f(x), f(w)) &\leq d_Y(f(x), f_L(x)) + d_Y(f_L(x), f_L(w)) \\ &\quad + d_Y(f_L(w), f(w)) \end{aligned}$$

for every $w \in X$, by the triangle inequality. It follows that

$$(1.8.11) \quad d_Y(f(x), f(w)) < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$$

for every $w \in X$ with $d_X(x, w) < \delta_L$, as desired.

1.8.3 Uniform convergence and uniform continuity

Similarly, if $\{f_j\}_{j=1}^\infty$ is a sequence of uniformly continuous mappings from X into Y that converges uniformly to a mapping f from X into Y , then

$$(1.8.12) \quad f \text{ is uniformly continuous on } X.$$

As before, we let $\epsilon > 0$ be given, and let L be a positive integer such that (1.8.8) holds for every $j \geq L$ and $w \in X$. In this case, the uniform continuity of f_L implies that there is a $\delta_L > 0$ such that (1.8.9) holds for every $x, w \in X$ with $d_X(x, w) < \delta_L$. This implies that (1.8.11) holds for every $x, w \in X$ with $d_X(x, w) < \delta_L$, as before.

1.8.4 Some limits of limits

Let E be a subset of X , and let $\{f_j\}_{j=1}^\infty$ be a sequence of mappings from E into Y that converges uniformly to a mapping f from E into Y . Suppose that $x \in X$ is a limit point of E . Suppose also that for each $j \geq 1$, the limit of $f_j(w)$ as $w \in E$ approaches x exists in Y , and put

$$(1.8.13) \quad q_j = \lim_{\substack{w \in E \\ w \rightarrow x}} f_j(w).$$

Under these conditions, one can show that

$$(1.8.14) \quad \{q_j\}_{j=1}^\infty \text{ is a Cauchy sequence in } Y.$$

Suppose in addition that

$$(1.8.15) \quad \{q_j\}_{j=1}^\infty \text{ converges to an element } q \text{ of } Y.$$

Of course, this holds automatically when Y is complete. In this case, one can show that

$$(1.8.16) \quad \lim_{\substack{w \in E \\ w \rightarrow x}} f(w) = q.$$

This corresponds to Theorem 7.11 on p149 of [210] when Y is the complex plane, with the standard metric. If Y is any complete metric space, then this corresponds to part of Exercise 17 on p168 of [210].

In particular, if x is a limit point of X , then we can take $E = X$, to get that f is continuous at x when f_j is continuous at x for each j , as before. In this case, $q_j = f_j(x)$ for each j , so that (1.8.15) holds with $q = f(x)$.

1.9 Bounded sets

Let (X, d_X) be a metric space. If $x \in X$ and r is a positive real number, then the *open ball* in X centered at x with radius r is defined as usual by

$$(1.9.1) \quad B(x, r) = B_X(x, r) = \{w \in X : d_X(x, w) < r\}.$$

If x' is another element of X , then it is easy to see that

$$(1.9.2) \quad B(x, r) \subseteq B(x', r + d_X(x, x')),$$

using the triangle inequality. It is well known that open balls in X are open sets.

A subset E of X is said to be *bounded* in X if there is an $x \in X$ and an $r > 0$ such that

$$(1.9.3) \quad E \subseteq B(x, r).$$

This implies that for every $x' \in X$ there is an $r' > 0$ such that

$$(1.9.4) \quad E \subseteq B(x', r'),$$

because of (1.9.2). Of course, this condition implies the previous one when $X \neq \emptyset$. To avoid minor technicalities, the empty set will be considered as a bounded set even when $X = \emptyset$.

1.9.1 Boundedness of compact sets

If K is a compact subset of X , then K is bounded in X . This is trivial when $X = \emptyset$, because the empty set is automatically considered to be a bounded set, and so we may suppose that $X \neq \emptyset$. If x is any element of X , then the collection of open balls $B(x, j)$ with $j \in \mathbf{Z}_+$ is an open covering of K , because

$$(1.9.5) \quad \bigcup_{j=1}^{\infty} B(x, j) = X.$$

If K is compact, then there are finitely many positive integers j_1, \dots, j_n such that

$$(1.9.6) \quad K \subseteq \bigcup_{l=1}^n B(x, j_l).$$

This implies that $K \subseteq B(x, r)$, with $r = \max(j_1, \dots, j_n)$.

1.9.2 More on bounded sets

Note that subsets of bounded sets are bounded. Let E_1, \dots, E_n be finitely many bounded subsets of X , and let us check that their union $\bigcup_{j=1}^n E_j$ is bounded. As before, this is trivial when $X = \emptyset$, and so we may suppose that $X \neq \emptyset$. If x is any element of X , then for each $j = 1, \dots, n$ there is a positive real number r_j such that

$$(1.9.7) \quad E_j \subseteq B(x, r_j).$$

This implies that

$$(1.9.8) \quad \bigcup_{j=1}^n E_j \subseteq B(x, \max_{1 \leq j \leq n} r_j),$$

as desired.

1.9.3 Closed balls

If $x \in X$ and r is a nonnegative real number, then the *closed ball* in X centered at x with radius r is defined by

$$(1.9.9) \quad \overline{B}(x, r) = \overline{B}_X(x, r) = \{w \in X : d_X(x, w) \leq r\}.$$

If x' is another element of X , then

$$(1.9.10) \quad \overline{B}(x, r) \subseteq \overline{B}(x', r + d_X(x, x')),$$

as in (1.9.2).

One can check that

$$(1.9.11) \quad \text{closed balls in } X \text{ are closed sets.}$$

One can verify that a closed ball contains all of its limit points, for instance, or that its complement in X is an open set.

Note that

$$(1.9.12) \quad B(x, r) \subseteq \overline{B}(x, r)$$

for every $r > 0$. If $t \in \mathbf{R}$ and $r < t$, then

$$(1.9.13) \quad \overline{B}(x, r) \subseteq B(x, t).$$

If $X \neq \emptyset$, then it is easy to see that a subset E of X is bounded if and only if it is contained in a closed ball in X . In this case, E is contained in a closed ball centered at any point in X , because of (1.9.10), as before. If E is bounded, then the closure \overline{E} of E in X is bounded too, because \overline{E} is contained in any closed ball in X that contains E .

1.9.4 Bounded sequences

A sequence $\{x_j\}_{j=1}^{\infty}$ of elements of X is said to be *bounded* in X if the set of x_j 's, $j \in \mathbf{Z}_+$, is bounded in X . It is well known and not difficult to check that

$$(1.9.14) \quad \text{convergent sequences in } X \text{ are bounded.}$$

Similarly, one can verify that

$$(1.9.15) \quad \text{Cauchy sequences in } X \text{ are bounded.}$$

Totally bounded subsets of metric spaces are discussed in Section 4.2. In particular, compact sets are totally bounded, and totally bounded sets are bounded.

1.10 Some remarks and examples

Let n be a positive integer, and let E be a subset of \mathbf{R}^n or \mathbf{C}^n . If E is bounded with respect to any of the metrics d_1 , d_2 , or d_∞ defined in Subsection 1.3.2 or Section 1.4, as appropriate, then it is easy to see that E is bounded with respect to the other two metrics, using the inequalities in Section 1.5. Similarly, if a sequence of elements of \mathbf{R}^n or \mathbf{C}^n converges with respect to any of these three metrics, then it converges with respect to the other two metrics, and with the same limit, as in Subsection 1.5.2.

Let X be a nonempty set, and let E be a subset of $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$. If E is bounded with respect to the metric d_2 defined in Subsection 1.6.1, then E is bounded with respect to d_∞ , because $d_\infty \leq d_2$, as in Subsection 1.6.2. Similarly, if E is bounded with respect to d_1 , then E is bounded with respect to d_2 and d_∞ . If X has only finitely many elements, and E is bounded with respect to d_∞ , then E is bounded with respect to d_1 and d_2 . This is essentially the same as in \mathbf{R}^n and \mathbf{C}^n when X has n elements.

Let $\{f_j\}_{j=1}^\infty$ be a sequence of elements of $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$, and let f be another element of the same space. If $\{f_j\}_{j=1}^\infty$ converges to f with respect to d_2 , then $\{f_j\}_{j=1}^\infty$ converges to f with respect to d_∞ , as mentioned in Subsection 1.6.2. Similarly, if $\{f_j\}_{j=1}^\infty$ converges to f with respect to d_1 , then $\{f_j\}_{j=1}^\infty$ converges to f with respect to d_2 and d_∞ . If X has only finitely many elements, and $\{f_j\}_{j=1}^\infty$ converges to f with respect to d_∞ , then $\{f_j\}_{j=1}^\infty$ converges to f with respect to d_1 and d_2 .

One can check that $\{f_j\}_{j=1}^\infty$ converges to f with respect to d_∞ if and only if $\{f_j\}_{j=1}^\infty$ converges to f uniformly on X . This uses the standard metric on \mathbf{R} or \mathbf{C} , as appropriate. A broader version of this will be discussed in Subsection 1.11.3.

1.10.1 A basis for $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$

If $x \in X$, then let δ_x be the real-valued function on X equal to 1 at x , and to 0 at every other element of X . It is easy to see that the collection of δ_x 's, $x \in X$, is a basis for each of $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$, as vector spaces over \mathbf{R} and \mathbf{C} , respectively, if one is familiar with these notions from linear algebra. Basically, this means that every element f of $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$ can be expressed in a unique way as a linear combination of the δ_x 's, $x \in X$, with coefficients in \mathbf{R} or \mathbf{C} , as appropriate. In fact, the coefficient of δ_x is equal to $f(x)$ for each $x \in X$. This means that f can be expressed as the sum of $f(x)\delta_x$ over the elements x of the support of f in X .

Observe that

$$(1.10.1) \quad \|\delta_x\|_1 = \|\delta_x\|_2 = \|\delta_x\|_\infty = 1$$

for every $x \in X$. If $x, y \in X$ and $x \neq y$, then

$$(1.10.2) \quad d_\infty(\delta_x, \delta_y) = \|\delta_x - \delta_y\|_\infty = 1,$$

$$(1.10.3) \quad d_2(\delta_x, \delta_y) = \|\delta_x - \delta_y\|_2 = \sqrt{2},$$

$$(1.10.4) \quad d_1(\delta_x, \delta_y) = \|\delta_x - \delta_y\|_1 = 2.$$

Let us now take $X = \mathbf{Z}_+$, and let δ_j be as before for each positive integer j . It is easy to see that $\{\delta_j\}_{j=1}^\infty$ converges to 0 pointwise on \mathbf{Z}_+ . Note that $\{\delta_j\}_{j=1}^\infty$ is bounded with respect to each of d_1 , d_2 , and d_∞ , and that $\{\delta_j\}_{j=1}^\infty$ does not converge to 0 with respect to any of these three metrics.

Similarly, $\{j\delta_j\}_{j=1}^\infty$ converges to 0 pointwise on \mathbf{Z}_+ . However,

$$(1.10.5) \quad \|j\delta_j\|_1 = \|j\delta_j\|_2 = \|j\delta_j\|_\infty = j$$

for each j , so that $\{j\delta_j\}_{j=1}^\infty$ is not bounded with respect to d_1 , d_2 , or d_∞ .

1.10.2 Some more examples in $c_{00}(\mathbf{Z}_+, \mathbf{R})$

If $j \in \mathbf{Z}_+$, then let f_j be the real-valued function on \mathbf{Z}_+ defined by

$$(1.10.6) \quad \begin{aligned} f_j(l) &= 1 && \text{when } l \leq j \\ &= 0 && \text{when } l > j. \end{aligned}$$

Observe that $\{f_j\}_{j=1}^\infty$ converges pointwise to the function equal to 1 everywhere on \mathbf{Z}_+ . One can check that $\{f_j\}_{j=1}^\infty$ does not converge uniformly on \mathbf{Z}_+ . We also have that

$$(1.10.7) \quad \|f_j\|_\infty = 1,$$

$$(1.10.8) \quad \|f_j\|_2 = \sqrt{j},$$

$$(1.10.9) \quad \|f_j\|_1 = j$$

for every j .

If $\alpha \in \mathbf{R}$, then

$$(1.10.10) \quad \|j^{-\alpha} f_j\|_\infty = j^{-\alpha} \|f_j\|_\infty = j^{-\alpha}$$

for every j . This means that $\{j^{-\alpha} f_j\}_{j=1}^\infty$ is bounded with respect to d_∞ exactly when $\alpha \geq 0$, and that $\{j^{-\alpha} f_j\}_{j=1}^\infty$ converges to 0 with respect to d_∞ exactly when $\alpha > 0$. Similarly,

$$(1.10.11) \quad \|j^{-\alpha} f_j\|_2 = j^{-\alpha} \|f_j\|_2 = j^{(1/2)-\alpha}$$

for every j , so that $\{j^{-\alpha} f_j\}_{j=1}^\infty$ is bounded with respect to d_2 if and only if $\alpha \geq 1/2$, and $\{j^{-\alpha} f_j\}_{j=1}^\infty$ converges to 0 with respect to d_2 exactly when $\alpha > 1/2$. In the same way,

$$(1.10.12) \quad \|j^{-\alpha} f_j\|_1 = j^{-\alpha} \|f_j\|_1 = j^{1-\alpha}$$

for every j , so that $\{j^{-\alpha} f_j\}_{j=1}^\infty$ is bounded with respect to d_1 if and only if $\alpha \geq 1$, and $\{j^{-\alpha} f_j\}_{j=1}^\infty$ converges to 0 with respect to d_1 if and only if $\alpha > 1$.

1.11 Bounded functions

Let X be a set, and let (Y, d_Y) be a metric space. A mapping f from X into Y is said to be *bounded* if

$$(1.11.1) \quad \text{the image } f(X) \text{ of } X \text{ under } f \text{ is a bounded subset of } Y.$$

Let $\mathcal{B}(X, Y)$ be the space of bounded mappings from X into Y .

1.11.1 Uniform convergence of bounded functions

Let $\{f_j\}_{j=1}^\infty$ be a sequence of bounded mappings from X into Y that converges uniformly to a mapping f from X into Y . This implies that there is an $L \in \mathbf{Z}_+$ such that

$$(1.11.2) \quad d_Y(f_j(x), f(x)) < 1$$

for every $j \geq L$ and $x \in X$. One can use this to check that

$$(1.11.3) \quad f \text{ is bounded,}$$

because f_L is bounded.

1.11.2 The supremum metric

Suppose that $X \neq \emptyset$, and let f, g be bounded mappings from X into Y . It is easy to see that

$$(1.11.4) \quad d_Y(f(x), g(x))$$

is bounded as a nonnegative real-valued function of x on X , using the triangle inequality. Put

$$(1.11.5) \quad \theta(f, g) = \sup\{d_Y(f(x), g(x)) : x \in X\}.$$

If $f = g$, then $f(x) = g(x)$ for every $x \in X$, so that

$$(1.11.6) \quad d_Y(f(x), g(x)) = 0$$

for every $x \in X$, and hence

$$(1.11.7) \quad \theta(f, g) = 0.$$

Conversely, if (1.11.7) holds, then (1.11.6) holds for every $x \in X$, so that $f(x) = g(x)$ for every $x \in X$, which means that $f = g$. We also have that

$$(1.11.8) \quad \theta(f, g) = \theta(g, f),$$

because $d_Y(f(x), g(x)) = d_Y(g(x), f(x))$ for every $x \in X$.

If h is another bounded mapping from X into Y , then

$$(1.11.9) \quad \begin{aligned} d_Y(f(x), h(x)) &\leq d_Y(f(x), g(x)) + d_Y(g(x), h(x)) \\ &\leq \theta(f, g) + \theta(g, h) \end{aligned}$$

for every $x \in X$. This implies that

$$(1.11.10) \quad \theta(f, h) \leq \theta(f, g) + \theta(g, h).$$

Thus (1.11.5) defines a metric on $\mathcal{B}(X, Y)$, which is known as the *supremum metric*.

1.11.3 Convergence with respect to the supremum metric

Let $\{f_j\}_{j=1}^{\infty}$ be a sequence of bounded mappings from X into Y , and let f be another bounded mapping from X into Y . If $\{f_j\}_{j=1}^{\infty}$ converges to f with respect to the supremum metric, then for each $\epsilon > 0$ there is an $L(\epsilon) \in \mathbf{Z}_+$ such that

$$(1.11.11) \quad \theta(f_j, f) < \epsilon$$

for every $j \geq L(\epsilon)$. It follows that

$$(1.11.12) \quad d_Y(f_j(x), f(x)) < \epsilon$$

for every $j \geq L(\epsilon)$ and $x \in X$, so that $\{f_j\}_{j=1}^{\infty}$ converges to f uniformly on X .

Conversely, if $\{f_j\}_{j=1}^{\infty}$ converges to f uniformly on X , then for each $\epsilon > 0$ there is an $L'(\epsilon) \in \mathbf{Z}_+$ such that (1.11.12) holds for every $j \geq L'(\epsilon)$ and $x \in X$. This implies that

$$(1.11.13) \quad \theta(f_j, f) = \sup\{d_Y(f_j(x), f(x)) : x \in X\} \leq \epsilon$$

for every $j \geq L'(\epsilon)$. One can use this to get that $\{f_j\}_{j=1}^{\infty}$ converges to f with respect to the supremum metric.

1.12 Completeness of $\mathcal{B}(X, Y)$

Let us continue with the same notation and hypotheses as in the previous section. Suppose that

$$(1.12.1) \quad Y \text{ is complete with respect to } d_Y,$$

and let us check that

$$(1.12.2) \quad \mathcal{B}(X, Y) \text{ is complete with respect to the supremum metric.}$$

1.12.1 The proof of completeness

Let $\{f_j\}_{j=1}^{\infty}$ be a sequence of bounded mappings from X into Y that is a Cauchy sequence with respect to the supremum metric. This means that for each $\epsilon > 0$ there is an $L(\epsilon) \in \mathbf{Z}_+$ such that

$$(1.12.3) \quad \theta(f_j, f_l) < \epsilon$$

for every $j, l \geq L(\epsilon)$. Thus

$$(1.12.4) \quad d_Y(f_j(x), f_l(x)) < \epsilon$$

for every $j, l \geq L(\epsilon)$ and $x \in X$.

In particular,

$$(1.12.5) \quad \{f_j(x)\}_{j=1}^{\infty} \text{ is a Cauchy sequence in } Y$$

for every $x \in X$. Because Y is complete,

$$(1.12.6) \quad \{f_j(x)\}_{j=1}^{\infty} \text{ converges in } Y$$

for every $x \in X$, and we put

$$(1.12.7) \quad f(x) = \lim_{j \rightarrow \infty} f_j(x).$$

This defines a mapping f from X into Y .

One can check that

$$(1.12.8) \quad d_Y(f_j(x), f(x)) \leq \epsilon$$

for every $j \geq L(\epsilon)$ and $x \in X$, using (1.12.4). Indeed,

$$(1.12.9) \quad \begin{aligned} d_Y(f_j(x), f(x)) &\leq d_Y(f_j(x), f_l(x)) + d_Y(f_l(x), f(x)) \\ &< \epsilon + d_Y(f_l(x), f(x)) \end{aligned}$$

for all $j, l \geq L(\epsilon)$ and $x \in X$, because of (1.12.4) and the triangle inequality. This implies (1.12.8), because $\{f_l(x)\}_{l=1}^{\infty}$ converges to $f(x)$ in Y .

It follows that $\{f_j\}_{j=1}^{\infty}$ converges to f uniformly on X . We also get that f is bounded on X , as in Subsection 1.11.1. This implies that $\{f_j\}_{j=1}^{\infty}$ converges to f with respect to the supremum metric, as in Subsection 1.11.3.

1.12.2 Necessity of completeness of Y

One can also verify that (1.12.2) implies (1.12.1). Indeed, every element of Y corresponds to a constant function on X with values in Y . It is easy to see that a Cauchy sequence in Y corresponds to a Cauchy sequence in $\mathcal{B}(X, Y)$ with respect to the supremum metric in this way. If a sequence of constant mappings from X to Y converges with respect to the supremum metric, then it converges uniformly on X , and thus pointwise on X , and it is easy to see that the limit is a constant mapping as well. This implies that the initial Cauchy sequence in Y converges to an element of Y .

Note that we get a mapping from Y into $\mathcal{B}(X, Y)$ in this way, by associating an element of Y with the corresponding constant function on X . This mapping is an *isometry*, in the sense that the metric on Y corresponds exactly to the distance between the corresponding constant functions on X with respect to the supremum metric. Isometric mappings between metric spaces are discussed further in Subsection 7.8.1.

1.13 More on bounded functions

The spaces of bounded real and complex-valued functions on a nonempty set X are also denoted

$$(1.13.1) \quad \ell^{\infty}(X, \mathbf{R}) \text{ and } \ell^{\infty}(X, \mathbf{C}),$$

respectively. This implicitly uses the standard Euclidean metrics on \mathbf{R} and \mathbf{C} .

If f and g are bounded real or complex-valued functions on X , then it is easy to see that

$$(1.13.2) \quad f + g \text{ is bounded on } X \text{ as well.}$$

Similarly, if f is a bounded real or complex-valued function on X , and t is a real or complex number, as appropriate, then

$$(1.13.3) \quad t f \text{ is bounded on } X$$

too. Thus $\ell^\infty(X, \mathbf{R})$ and $\ell^\infty(X, \mathbf{C})$ are linear subspaces of the real and complex vector spaces of all real and complex-valued functions on X , respectively.

1.13.1 The supremum norm

If f is a bounded real or complex-valued function on X , then put

$$(1.13.4) \quad \|f\|_\infty = \sup\{|f(x)| : x \in X\}.$$

Note that $\|f\|_\infty = 0$ if and only if $f \equiv 0$ on X . If $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then one can check that

$$(1.13.5) \quad \|t f\|_\infty = |t| \|f\|_\infty.$$

If g is another bounded real or complex-valued function on X , then

$$(1.13.6) \quad |f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty$$

for every $x \in X$. This implies that

$$(1.13.7) \quad \|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty.$$

It follows that (1.13.4) defines a norm on each of $\ell^\infty(X, \mathbf{R})$ and $\ell^\infty(X, \mathbf{C})$, which is known as the *supremum norm*. The corresponding metric

$$(1.13.8) \quad d_\infty(f, g) = \|f - g\|_\infty = \sup\{|f(x) - g(x)| : x \in X\}$$

is the same as the supremum metric on these spaces, associated to the standard Euclidean metric on \mathbf{R} or \mathbf{C} , as appropriate.

Note that $\ell^\infty(\mathbf{Z}_+, \mathbf{R})$ is often simply denoted ℓ^∞ . This corresponds to the example in 4.2 C 4 on p107 of [93].

1.14 Continuous functions

Let (X, d_X) and (Y, d_Y) be metric spaces, and let $C(X, Y)$ be the space of continuous mappings from X into Y . Also let

$$(1.14.1) \quad C_b(X, Y) = \mathcal{B}(X, Y) \cap C(X, Y)$$

be the space of bounded continuous mappings from X into Y . If f is a continuous mapping from X into Y and X is compact, then it is well known that $f(X)$ is compact in Y , so that $f(X)$ is bounded in Y in particular. Thus $C_b(X, Y)$ is the same as $C(X, Y)$ when X is compact.

1.14.1 The supremum metric and $C_b(X, Y)$

Suppose that $X \neq \emptyset$, so that the supremum metric can be defined on $\mathcal{B}(X, Y)$ as in Subsection 1.11.2. Note that

$$(1.14.2) \quad C_b(X, Y) \text{ is a closed set in } \mathcal{B}(X, Y),$$

with respect to the supremum metric. More precisely, if $\{f_j\}_{j=1}^\infty$ is a sequence of bounded continuous mappings from X into Y that converges to a bounded mapping f from X into Y with respect to the supremum metric, then we have seen that $\{f_j\}_{j=1}^\infty$ converges to f uniformly on X , and hence that f is continuous on X . Alternatively, if $f \in \mathcal{B}(X, Y)$ is a limit point of $C_b(X, Y)$ with respect to the supremum metric, then one can show that f is continuous on X . This is very similar to the argument used to show that uniform limits of continuous mappings are continuous, as in Subsection 1.8.2.

1.14.2 Completeness of $C_b(X, Y)$

Of course, $C_b(X, Y)$ may be considered as a metric space, using the restriction of the supremum metric on $\mathcal{B}(X, Y)$ to $C_b(X, Y)$. Suppose that Y is complete with respect to d_Y , so that $\mathcal{B}(X, Y)$ is complete with respect to the supremum metric, as in Section 1.12. Under these conditions, we get that

$$(1.14.3) \quad C_b(X, Y) \text{ is complete as a metric space}$$

with respect to the supremum metric, as in Subsection 1.7.1. This also uses (1.14.2). One could show this more directly using an argument like that in Subsection 1.12.1 as well.

Note that the completeness of Y is necessary to have (1.14.3), for essentially the same reasons as in Subsection 1.12.2. This uses the fact that constant functions on X are continuous. In particular, the mapping from Y into $\mathcal{B}(X, Y)$ mentioned earlier in fact maps Y into $C_b(X, Y)$.

1.14.3 Real and complex-valued functions

Let us now take $Y = \mathbf{R}$ or \mathbf{C} , with their standard Euclidean metrics. If f and g are continuous real or complex-valued functions on X , then it is well known that their sum $f + g$ is continuous on X too. Similarly, if f is a continuous real or complex-valued function on X , and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then tf is continuous on X . This means that $C(X, \mathbf{R})$ and $C(X, \mathbf{C})$ are linear subspaces of the real and complex vector spaces of all real and complex-valued functions on X , respectively. We may also consider

$$(1.14.4) \quad C_b(X, \mathbf{R}) = \ell^\infty(X, \mathbf{R}) \cap C(X, \mathbf{R})$$

as a linear subspace of both $\ell^\infty(X, \mathbf{R})$ and $C(X, \mathbf{R})$, and

$$(1.14.5) \quad C_b(X, \mathbf{C}) = \ell^\infty(X, \mathbf{C}) \cap C(X, \mathbf{C})$$

as a linear subspace of both $\ell^\infty(X, \mathbf{C})$ and $C(X, \mathbf{C})$.

1.15 Continuous functions on $[0, 1]$

In this section, we take X to be the closed unit interval $[0, 1]$ in the real line, equipped with the restriction of the standard Euclidean metric on \mathbf{R} to $[0, 1]$. It is well known that $[0, 1]$ is compact as a subset of \mathbf{R} , and thus as a subset of itself. This means that every continuous real or complex-valued function f on $[0, 1]$ is bounded, as before.

1.15.1 Norms on $C([0, 1], \mathbf{R})$, $C([0, 1], \mathbf{C})$

A nonnegative real-valued function N on $C([0, 1], \mathbf{R})$ or $C([0, 1], \mathbf{C})$ is said to be a *norm* if it satisfies the usual three conditions, as follows. First, $N(f) = 0$ if and only if $f = 0$. Second, if f is a continuous real or complex-valued function on $[0, 1]$ and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then

$$(1.15.1) \quad N(tf) = |t|N(f).$$

Third, if f and g are continuous real or complex-valued functions on $[0, 1]$, then

$$(1.15.2) \quad N(f + g) \leq N(f) + N(g).$$

In this case,

$$(1.15.3) \quad d_N(f, g) = N(f - g)$$

defines a metric on $C([0, 1], \mathbf{R})$ or $C([0, 1], \mathbf{C})$, as appropriate. The supremum norm (1.13.4) defines a norm on each of $C([0, 1], \mathbf{R})$ and $C([0, 1], \mathbf{C})$, for which the corresponding metric (1.13.8) is the supremum metric, as in Subsection 1.13.1.

1.15.2 The L^1 norm on $C([0, 1], \mathbf{R})$ and $C([0, 1], \mathbf{C})$

If f is a continuous real or complex-valued function on $[0, 1]$, then put

$$(1.15.4) \quad \|f\|_1 = \int_0^1 |f(x)| dx.$$

More precisely, it is well known and not difficult to verify that $|f(x)|$ is also continuous on $[0, 1]$, so that the Riemann integral on the right side of (1.15.4) exists. If $f(x_0) \neq 0$ for some $0 \leq x_0 \leq 1$, then

$$(1.15.5) \quad |f(x)| \geq |f(x_0)|/2 > 0$$

when $0 \leq x \leq 1$ is sufficiently close to x_0 , because f is continuous at x_0 . This implies that $\|f\|_1 > 0$, so that (1.15.4) satisfies the first condition in the definition of a norm. It is easy to see that (1.15.4) satisfies (1.15.1) and (1.15.2), so that (1.15.4) defines a norm on $C([0, 1], \mathbf{R})$ and $C([0, 1], \mathbf{C})$.

1.15.3 The L^2 norm on $C([0, 1], \mathbf{R})$ and $C([0, 1], \mathbf{C})$

Similarly, put

$$(1.15.6) \quad \|f\|_2 = \left(\int_0^1 |f(x)|^2 dx \right)^{1/2}$$

for every continuous real or complex-valued function f on $[0, 1]$. One can check that (1.15.6) is equal to 0 exactly when $f = 0$ on $[0, 1]$, using the same type of argument as in the preceding paragraph. It is easy to see that (1.15.6) satisfies the homogeneity condition (1.15.1). If g is another continuous real or complex-valued function on $[0, 1]$, then it is well known that

$$(1.15.7) \quad \|fg\|_1 \leq \|f\|_2 \|g\|_2,$$

which is an integral version of the Cauchy–Schwarz inequality. A proof of this may be found in Subsection 3.2.1. This can be used to show that (1.15.6) satisfies the triangle inequality (1.15.2), as in Subsection 3.2.2. It follows that (1.15.6) defines a norm on $C([0, 1], \mathbf{R})$ and $C([0, 1], \mathbf{C})$.

1.15.4 More on these integral norms

It is easy to see that

$$(1.15.8) \quad \|f\|_1, \|f\|_2 \leq \|f\|_\infty$$

for every continuous real or complex-valued function f on $[0, 1]$. One can also get that

$$(1.15.9) \quad \|f\|_1 \leq \|f\|_2,$$

using (1.15.7). Let

$$(1.15.10) \quad d_1(f, g) = \|f - g\|_1$$

and

$$(1.15.11) \quad d_2(f, g) = \|f - g\|_2$$

be the metrics on $C([0, 1], \mathbf{R})$ and $C([0, 1], \mathbf{C})$ associated to (1.15.4) and (1.15.6) as in (1.15.3), respectively. Using (1.15.8) and (1.15.9), we get that

$$(1.15.12) \quad d_1(f, g) \leq d_2(f, g) \leq d_\infty(f, g)$$

for all continuous real and complex-valued functions f and g on $[0, 1]$.

It follows that convergent sequences in $C([0, 1], \mathbf{R})$ and $C([0, 1], \mathbf{C})$ with respect to d_∞ are convergent with respect to d_1 and d_2 , with the same limit. Similarly, convergent sequences with respect to d_2 are convergent with respect to d_1 , with the same limit. Some examples of sequences of continuous real and complex-valued functions on $[0, 1]$ are discussed in Section A.4.

We also have that Cauchy sequences in $C([0, 1], \mathbf{R})$ and $C([0, 1], \mathbf{C})$ with respect to d_2 are Cauchy sequences with respect to d_1 . Some examples of Cauchy sequences with respect to d_1 and d_2 are discussed in Section A.5. In particular, neither $C([0, 1], \mathbf{R})$ nor $C([0, 1], \mathbf{C})$ is complete with respect to either d_1 or d_2 . To get complete metric spaces, one can use *Lebesgue integrals*.

Of course, bounded sets and sequences in $C([0, 1], \mathbf{R})$ and $C([0, 1], \mathbf{C})$ with respect to d_∞ are bounded with respect to d_1 and d_2 as well. Similarly, boundedness with respect to d_2 implies boundedness with respect to d_1 . This is related to the examples in Sections A.4 and A.5 too.

Note that the identity mappings on $C([0, 1], \mathbf{R})$ and $C([0, 1], \mathbf{C})$ are Lipschitz with respect to d_∞ on the domain and d_1 or d_2 on the range. Similarly, the identity mappings on $C([0, 1], \mathbf{R})$ and $C([0, 1], \mathbf{C})$ are Lipschitz with respect to d_2 on the domain and d_1 on the range. Other properties of these metrics can be obtained from this, as before.

Chapter 2

Basic ℓ^1 and ℓ^2 spaces

This chapter deals with classical ℓ^1 and ℓ^2 spaces, of absolutely summable and square-summable sequences of real or complex numbers, respectively. We also consider c_0 spaces of real or complex-valued functions on arbitrary nonempty sets that vanish at infinity. Sums over arbitrary nonempty sets and corresponding ℓ^1 , ℓ^2 spaces will be discussed in Chapter 11.

2.1 Infinite series

Remember that an infinite series

$$(2.1.1) \quad \sum_{j=1}^{\infty} a_j$$

of real or complex numbers is said to *converge* if the corresponding sequence of partial sums

$$(2.1.2) \quad \sum_{j=1}^n a_j$$

converges with respect to the standard Euclidean metric on \mathbf{R} or \mathbf{C} , as appropriate. Of course, the value of the sum (2.1.1) is defined to be the limit of the sequence of partial sums (2.1.2) in this case.

If (2.1.1) converges, and if $\sum_{j=1}^{\infty} b_j$ is another convergent series of real or complex numbers, as appropriate, then $\sum_{j=1}^{\infty} (a_j + b_j)$ converges, with

$$(2.1.3) \quad \sum_{j=1}^{\infty} (a_j + b_j) = \sum_{j=1}^{\infty} a_j + \sum_{j=1}^{\infty} b_j.$$

This reduces to the corresponding statement for sums of convergent sequences of real or complex numbers, applied to the partial sums of these series. Similarly,

if (2.1.1) converges, and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then $\sum_{j=1}^{\infty} t a_j$ converges, with

$$(2.1.4) \quad \sum_{j=1}^{\infty} t a_j = t \sum_{j=1}^{\infty} a_j.$$

2.1.1 Series with nonnegative terms

If a_j is a nonnegative real number for each $j \geq 1$, then the partial sums (2.1.2) increase monotonically. It is well known that a monotonically increasing sequence of real numbers converges with respect to the standard Euclidean metric on \mathbf{R} if and only if the sequence has an upper bound in \mathbf{R} . In this case, the sequence converges to the supremum of the set of its terms.

If a_j is a nonnegative real number for each j , but the partial sums (2.1.2) do not have an upper bound in \mathbf{R} , then it is sometimes convenient to consider the value of the sum (2.1.1) to be $+\infty$. Note that the partial sums (2.1.2) tend to $+\infty$ as $n \rightarrow \infty$ in this situation.

If a_j and b_j are nonnegative real numbers for each $j \geq 1$, then one can check that (2.1.3) holds, where the right side of (2.1.3) is considered to be $+\infty$ when either of the individual sums is $+\infty$. Similarly, (2.1.4) holds for every positive real number t , where the right side is considered to be $+\infty$ when (2.1.1) is $+\infty$.

If a_j and b_j are nonnegative real numbers with

$$(2.1.5) \quad a_j \leq b_j$$

for every $j \geq 1$, then

$$(2.1.6) \quad \sum_{j=1}^{\infty} a_j \leq \sum_{j=1}^{\infty} b_j,$$

which is trivial when the right side is $+\infty$.

2.1.2 Absolute convergence

An infinite series (2.1.1) of real or complex numbers is said to converge *absolutely* if

$$(2.1.7) \quad \sum_{j=1}^{\infty} |a_j|$$

converges as an infinite series of nonnegative real numbers. This means that (2.1.7) is finite, in terms of the conventions for sums of nonnegative real numbers mentioned in the previous subsection.

If (2.1.1) converges absolutely, then it is well known that (2.1.1) converges in the usual sense. One can also check that

$$(2.1.8) \quad \left| \sum_{j=1}^{\infty} a_j \right| \leq \sum_{j=1}^{\infty} |a_j|$$

under these conditions. This uses the fact that

$$(2.1.9) \quad \left| \sum_{j=1}^n a_j \right| \leq \sum_{j=1}^n |a_j|$$

for every positive integer n , by the triangle inequality.

If (2.1.1) converges, then it is well known that $\{a_j\}_{j=1}^\infty$ converges to 0 as a sequence of real or complex numbers, as appropriate. In particular, this holds when (2.1.1) converges absolutely.

2.2 Basic ℓ^1 spaces

Let $\ell^1(\mathbf{Z}_+, \mathbf{R})$ be the space of real-valued functions f on the set \mathbf{Z}_+ of positive integers such that

$$(2.2.1) \quad \|f\|_1 = \sum_{j=1}^{\infty} |f(j)|$$

is finite, which is to say that the right side converges as an infinite series of nonnegative real numbers. This is also often simply denoted ℓ^1 . Similarly, let $\ell^1(\mathbf{Z}_+, \mathbf{C})$ be the space of complex-valued functions f on \mathbf{Z}_+ such that (2.2.1) is finite. In both cases, the convergence of the series on the right side of (2.2.1) implies that

$$(2.2.2) \quad \lim_{j \rightarrow \infty} f(j) = 0,$$

as mentioned in the previous section.

If $f \in \ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, then (2.2.1) is a nonnegative real number, which is equal to 0 exactly when $f(j) = 0$ for every $j \in \mathbf{Z}_+$. If $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then $t f(j)$ defines another real or complex-valued function on \mathbf{Z}_+ , as appropriate, and

$$(2.2.3) \quad \|t f\|_1 = \sum_{j=1}^{\infty} |t f(j)| = |t| \|f\|_1.$$

In particular, $t f \in \ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, as appropriate.

Let g be another element of $\ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, as appropriate. Thus $f + g$ is a real or complex-valued function on \mathbf{Z}_+ , as appropriate, and

$$(2.2.4) \quad \|f + g\|_1 = \sum_{j=1}^{\infty} |f(j) + g(j)| \leq \sum_{j=1}^{\infty} (|f(j)| + |g(j)|) = \|f\|_1 + \|g\|_1.$$

This implies that $f + g \in \ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, as appropriate. Hence $\ell^1(\mathbf{Z}_+, \mathbf{R})$ and $\ell^1(\mathbf{Z}_+, \mathbf{C})$ are linear subspaces of the real and complex vector spaces of all real and complex-valued functions on \mathbf{Z}_+ , respectively.

This means that $\ell^1(\mathbf{Z}_+, \mathbf{R})$ and $\ell^1(\mathbf{Z}_+, \mathbf{C})$ are vector spaces over the real and complex numbers, respectively, with respect to pointwise addition and scalar multiplication. We also get that (2.2.1) defines a norm on each of these spaces,

by the remarks in the preceding paragraphs. Using (2.2.3) and (2.2.4), one can check that

$$(2.2.5) \quad d_1(f, g) = \|f - g\|_1$$

defines a metric on each of $\ell^1(\mathbf{Z}_+, \mathbf{R})$ and $\ell^1(\mathbf{Z}_+, \mathbf{C})$, as usual.

2.2.1 More on $\ell^1(\mathbf{Z}_+, \mathbf{R})$, $\ell^1(\mathbf{Z}_+, \mathbf{C})$

If $f \in \ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, then

$$(2.2.6) \quad \sum_{j=1}^{\infty} f(j)$$

converges as an infinite series of real or complex numbers, as in the previous section. If g is another element of $\ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, as appropriate, then

$$(2.2.7) \quad \left| \sum_{j=1}^{\infty} f(j) - \sum_{j=1}^{\infty} g(j) \right| = \left| \sum_{j=1}^{\infty} (f(j) - g(j)) \right| \\ \leq \sum_{j=1}^{\infty} |f(j) - g(j)| = \|f - g\|_1,$$

using (2.1.8) in the second step. This implies that the mapping from f to the sum (2.2.6) is Lipschitz with constant $C = 1$ as a mapping from $\ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$ into \mathbf{R} or \mathbf{C} , respectively, using the ℓ^1 metric (2.2.5) on $\ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, and the standard Euclidean metric on \mathbf{R} or \mathbf{C} .

If $f \in \ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, then it is easy to see directly that f is bounded on \mathbf{Z}_+ , with

$$(2.2.8) \quad \|f\|_{\infty} \leq \|f\|_1.$$

Here $\|f\|_{\infty}$ is the supremum norm of f on \mathbf{Z}_+ , as in (1.13.4). If g is another element of $\ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, as appropriate, then we get that

$$(2.2.9) \quad d_{\infty}(f, g) \leq d_1(f, g),$$

where $d_{\infty}(f, g)$ is the supremum metric for bounded real or complex-valued functions on \mathbf{Z}_+ , as in (1.13.8).

Remember that $c_{00}(\mathbf{Z}_+, \mathbf{R})$, $c_{00}(\mathbf{Z}_+, \mathbf{C})$ are the spaces of real and complex-valued functions on \mathbf{Z}_+ with finite support, respectively, as in Section 1.6. Clearly

$$(2.2.10) \quad c_{00}(\mathbf{Z}_+, \mathbf{R}) \subseteq \ell^1(\mathbf{Z}_+, \mathbf{R}), \quad c_{00}(\mathbf{Z}_+, \mathbf{C}) \subseteq \ell^1(\mathbf{Z}_+, \mathbf{C}),$$

since an infinite series automatically converges when all but finitely many terms are equal to 0. One can also check that $c_{00}(\mathbf{Z}_+, \mathbf{R})$ and $c_{00}(\mathbf{Z}_+, \mathbf{C})$ are dense subsets of $\ell^1(\mathbf{Z}_+, \mathbf{R})$ and $\ell^1(\mathbf{Z}_+, \mathbf{C})$, respectively, with respect to the ℓ^1 metric (2.2.5).

2.3 Basic ℓ^2 spaces

Let $\ell^2(\mathbf{Z}_+, \mathbf{R})$ and $\ell^2(\mathbf{Z}_+, \mathbf{C})$ be the spaces of real and complex-valued functions f on \mathbf{Z}_+ such that

$$(2.3.1) \quad \sum_{j=1}^{\infty} |f(j)|^2$$

converges as an infinite series of nonnegative real numbers, respectively. In both cases, we put

$$(2.3.2) \quad \|f\|_2 = \left(\sum_{j=1}^{\infty} |f(j)|^2 \right)^{1/2},$$

using the nonnegative square root on the right side. Note that this is equal to 0 if and only if $f(j) = 0$ for every $j \geq 1$.

As before, $\ell^2(\mathbf{Z}_+, \mathbf{R})$ is often simply denoted ℓ^2 . This is discussed in Section 3.10 of [93], and in the example in 4.2 C 5 on p107 of [93]. This corresponds to the infinite-dimensional Euclidean space discussed in Section 8 of Chapter 2 of [177] as well.

The convergence of the series (2.3.1) implies that

$$(2.3.3) \quad \lim_{j \rightarrow \infty} |f(j)|^2 = 0,$$

as before. It follows that $f(j) \rightarrow 0$ as $j \rightarrow \infty$.

Let $f \in \ell^2(\mathbf{Z}_+, \mathbf{R})$ or $\ell^2(\mathbf{Z}_+, \mathbf{C})$ be given. If $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then

$$(2.3.4) \quad \sum_{j=1}^{\infty} |t f(j)|^2 = |t|^2 \sum_{j=1}^{\infty} |f(j)|^2,$$

and in particular the series on the left converges. This means that $t f$ is an element of $\ell^2(\mathbf{Z}_+, \mathbf{R})$ or $\ell^2(\mathbf{Z}_+, \mathbf{C})$, as appropriate, with

$$(2.3.5) \quad \|t f\|_2 = |t| \|f\|_2.$$

2.3.1 The Cauchy–Schwarz inequality for infinite series

If a and b are nonnegative real numbers, then it is well known that

$$(2.3.6) \quad a b \leq \frac{1}{2} (a^2 + b^2),$$

because $0 \leq (a - b)^2 = a^2 - 2 a b + b^2$. Suppose that g is another element of $\ell^2(\mathbf{Z}_+, \mathbf{R})$ or $\ell^2(\mathbf{Z}_+, \mathbf{C})$, as appropriate. Using (2.3.6), we get that

$$(2.3.7) \quad |f(j)| |g(j)| \leq \frac{1}{2} (|f(j)|^2 + |g(j)|^2)$$

for every $j \in \mathbf{Z}_+$. Thus

$$\begin{aligned}
 \sum_{j=1}^{\infty} |f(j)| |g(j)| &\leq \sum_{j=1}^{\infty} \frac{1}{2} (|f(j)|^2 + |g(j)|^2) \\
 (2.3.8) \qquad &= \frac{1}{2} \sum_{j=1}^{\infty} |f(j)|^2 + \frac{1}{2} \sum_{j=1}^{\infty} |g(j)|^2 = \frac{1}{2} \|f\|_2^2 + \frac{1}{2} \|g\|_2^2.
 \end{aligned}$$

In particular, the series on the left converges, so that $f g \in \ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, as appropriate.

In fact, we have that

$$(2.3.9) \qquad \sum_{j=1}^{\infty} |f(j)| |g(j)| \leq \|f\|_2 \|g\|_2$$

under these conditions, which is another version of the Cauchy–Schwarz inequality. This follows from (2.3.8) when $\|f\|_2 = \|g\|_2 = 1$, and otherwise one can reduce to that case using (2.3.5).

2.3.2 The triangle inequality for the ℓ^2 norm

Observe that

$$(2.3.10) \quad |f(j) + g(j)|^2 \leq (|f(j)| + |g(j)|)^2 = |f(j)|^2 + 2|f(j)| |g(j)| + |g(j)|^2$$

for every $j \geq 1$, so that

$$(2.3.11) \quad \sum_{j=1}^{\infty} |f(j) + g(j)|^2 \leq \sum_{j=1}^{\infty} |f(j)|^2 + 2 \sum_{j=1}^{\infty} |f(j)| |g(j)| + \sum_{j=1}^{\infty} |g(j)|^2.$$

This implies that the series on the left converges, so that $f + g \in \ell^2(\mathbf{Z}_+, \mathbf{R})$ or $\ell^2(\mathbf{Z}_+, \mathbf{C})$, as appropriate. Combining (2.3.9) and (2.3.11), we get that

$$(2.3.12) \quad \|f + g\|_2^2 \leq \|f\|_2^2 + 2\|f\|_2 \|g\|_2 + \|g\|_2^2 = (\|f\|_2 + \|g\|_2)^2.$$

This means that

$$(2.3.13) \qquad \|f + g\|_2 \leq \|f\|_2 + \|g\|_2.$$

This shows that $\ell^2(\mathbf{Z}_+, \mathbf{R})$ and $\ell^2(\mathbf{Z}_+, \mathbf{C})$ are linear subspaces of the real and complex vector spaces of all real and complex-valued functions on \mathbf{Z}_+ , respectively. We also get that (2.3.2) defines a norm on each of $\ell^2(\mathbf{Z}_+, \mathbf{R})$ and $\ell^2(\mathbf{Z}_+, \mathbf{C})$, by (2.3.5) and (2.3.13). It follows that

$$(2.3.14) \qquad d_2(f, g) = \|f - g\|_2$$

defines a metric on each of $\ell^2(\mathbf{Z}_+, \mathbf{R})$ and $\ell^2(\mathbf{Z}_+, \mathbf{C})$, as usual.

2.3.3 More on $\ell^2(\mathbf{Z}_+, \mathbf{R})$, $\ell^2(\mathbf{Z}_+, \mathbf{C})$

If $f \in \ell^2(\mathbf{Z}_+, \mathbf{R})$ or $\ell^2(\mathbf{Z}_+, \mathbf{C})$, then one can check directly that f is bounded on \mathbf{Z}_+ , with

$$(2.3.15) \quad \|f\|_\infty \leq \|f\|_2.$$

If g is another element of $\ell^2(\mathbf{Z}_+, \mathbf{R})$ or $\ell^2(\mathbf{Z}_+, \mathbf{C})$, as appropriate, then we get that

$$(2.3.16) \quad d_\infty(f, g) \leq d_2(f, g).$$

Suppose now that $f \in \ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$. Observe that

$$(2.3.17) \quad \sum_{j=1}^{\infty} |f(j)|^2 \leq \|f\|_\infty \sum_{j=1}^{\infty} |f(j)| = \|f\|_\infty \|f\|_1 \leq \|f\|_1^2,$$

using (2.2.8) in the third step. This implies that $f \in \ell^2(\mathbf{Z}_+, \mathbf{R})$ or $\ell^2(\mathbf{Z}_+, \mathbf{C})$, as appropriate, with

$$(2.3.18) \quad \|f\|_2 \leq \|f\|_1.$$

In particular, we get that

$$(2.3.19) \quad \ell^1(\mathbf{Z}_+, \mathbf{R}) \subseteq \ell^2(\mathbf{Z}_+, \mathbf{R}), \quad \ell^1(\mathbf{Z}_+, \mathbf{C}) \subseteq \ell^2(\mathbf{Z}_+, \mathbf{C}).$$

If g is another element of $\ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, as appropriate, then it follows that

$$(2.3.20) \quad d_2(f, g) \leq d_1(f, g).$$

If f is a real or complex-valued function on \mathbf{Z}_+ with finite support, then f is clearly an element of $\ell^2(\mathbf{Z}_+, \mathbf{R})$ or $\ell^2(\mathbf{Z}_+, \mathbf{C})$, as appropriate. One can verify that $c_{00}(\mathbf{Z}_+, \mathbf{R})$ and $c_{00}(\mathbf{Z}_+, \mathbf{C})$ are dense subsets of $\ell^2(\mathbf{Z}_+, \mathbf{R})$ and $\ell^2(\mathbf{Z}_+, \mathbf{C})$, respectively, with respect to the ℓ^2 metric (2.3.14).

2.4 Completeness of ℓ^1 , ℓ^2

Let $\{f_l\}_{l=1}^\infty$ be a sequence of elements of $\ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, and suppose that the f_l 's have bounded ℓ^1 norms, so that there is a nonnegative real number C_1 such that

$$(2.4.1) \quad \sum_{j=1}^{\infty} |f_l(j)| \leq C_1$$

for every $l \geq 1$. Suppose also that $\{f_l\}_{l=1}^\infty$ converges to a real or complex-valued function f pointwise on \mathbf{Z}_+ , as appropriate. Let us verify that $f \in \ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, as appropriate, with

$$(2.4.2) \quad \sum_{j=1}^{\infty} |f(j)| \leq C_1.$$

If $n \in \mathbf{Z}_+$, then

$$(2.4.3) \quad \sum_{j=1}^n |f(j)| = \lim_{l \rightarrow \infty} \sum_{j=1}^n |f_l(j)| \leq C_1,$$

using pointwise convergence in the first step, and (2.4.1) in the second step. This implies (2.4.2), since this estimate holds for all $n \geq 1$.

2.4.1 Completeness of ℓ^1

We would like to show that $\ell^1(\mathbf{Z}_+, \mathbf{R})$ and $\ell^1(\mathbf{Z}_+, \mathbf{C})$ are complete with respect to the ℓ^1 metric (2.2.5). Let $\{f_l\}_{l=1}^\infty$ be a sequence of elements of $\ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$ that is a Cauchy sequence with respect to (2.2.5). This implies that for every $\epsilon > 0$ there is a positive integer $L(\epsilon)$ such that

$$(2.4.4) \quad \sum_{j=1}^\infty |f_k(j) - f_l(j)| = \|f_k - f_l\|_1 < \epsilon$$

for all $k, l \geq L(\epsilon)$. In particular,,

$$(2.4.5) \quad |f_k(j) - f_l(j)| < \epsilon$$

for every $j \in \mathbf{Z}_+$ when $k, l \geq L(\epsilon)$.

This means that $\{f_l(j)\}_{l=1}^\infty$ is a Cauchy sequence of real or complex numbers, as appropriate, for every $j \in \mathbf{Z}_+$, and with respect to the standard Euclidean metric on \mathbf{R} or \mathbf{C} . Remember that \mathbf{R} and \mathbf{C} are complete as metric spaces with respect to their standard Euclidean metrics. It follows that $\{f_l(j)\}_{l=1}^\infty$ converges in \mathbf{R} or \mathbf{C} , as appropriate, for every $j \in \mathbf{Z}_+$, with respect to the standard Euclidean metric.

Thus

$$(2.4.6) \quad f(j) = \lim_{l \rightarrow \infty} f_l(j)$$

defines a real or complex-valued function on \mathbf{Z}_+ , as appropriate. We would like to check that $\sum_{j=1}^\infty |f(j)|$ converges, so that f is an element of $\ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, as appropriate. We would also like to verify that $\{f_l\}_{l=1}^\infty$ converges to f with respect to the ℓ^1 metric.

Let $\epsilon > 0$ and $l \geq L(\epsilon)$ be given. Note that $\{f_k - f_l\}_{k=L(\epsilon)}^\infty$ is a sequence of elements of $\ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, as appropriate, that converges to $f - f_l$ pointwise on \mathbf{Z}_+ . The remarks at the beginning of the section imply that $f - f_l$ is an element of $\ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, as appropriate, with

$$(2.4.7) \quad \|f - f_l\|_1 = \sum_{j=1}^\infty |f(j) - f_l(j)| \leq \epsilon,$$

because of (2.4.4). Hence $f \in \ell^1(\mathbf{Z}_+, \mathbf{R})$ or $\ell^1(\mathbf{Z}_+, \mathbf{C})$, as appropriate, because of the corresponding property of f_l . It is easy to see that $\{f_l\}_{l=1}^\infty$ converges to f with respect to the ℓ^1 metric, because (2.4.7) holds for every $l \geq L(\epsilon)$.

2.4.2 Completeness of ℓ^2

Now let $\{f_l\}_{l=1}^\infty$ be a sequence of elements of $\ell^2(\mathbf{Z}_+, \mathbf{R})$ or $\ell^2(\mathbf{Z}_+, \mathbf{C})$ with bounded ℓ^2 norms, so that

$$(2.4.8) \quad \|f_l\|_2 \leq C_2$$

for some $C_2 \geq 0$ and every $l \geq 1$. This is the same as saying that

$$(2.4.9) \quad \sum_{j=1}^{\infty} |f_l(j)|^2 \leq C_2^2$$

for every $l \geq 1$. Suppose that $\{f_l\}_{l=1}^\infty$ also converges pointwise to a real or complex-valued function f on \mathbf{Z}_+ , which implies that $\{|f_l|^2\}_{l=1}^\infty$ converges to $|f|^2$ pointwise on \mathbf{Z}_+ too. It follows that

$$(2.4.10) \quad \sum_{j=1}^{\infty} |f(j)|^2 \leq C_2^2,$$

by the remarks at the beginning of the section, applied to $\{|f_l|^2\}_{l=1}^\infty$. This means that $f \in \ell^2(\mathbf{Z}_+, \mathbf{R})$ or $\ell^2(\mathbf{Z}_+, \mathbf{C})$, as appropriate, with

$$(2.4.11) \quad \|f\|_2 \leq C_2.$$

Using this, one can show that $\ell^2(\mathbf{Z}_+, \mathbf{R})$ and $\ell^2(\mathbf{Z}_+, \mathbf{C})$ are complete with respect to the ℓ^2 metric (2.3.14). The argument is similar to the previous one for $\ell^1(\mathbf{Z}_+, \mathbf{R})$, $\ell^1(\mathbf{Z}_+, \mathbf{C})$.

2.5 Vanishing at infinity

Let X be a nonempty set, and let f be a real or complex-valued function on X . We say that f *vanishes at infinity* on X if for every $\epsilon > 0$,

$$(2.5.1) \quad |f(x)| < \epsilon$$

for all but finitely many $x \in X$. Equivalently, this means that for each $\epsilon > 0$,

$$(2.5.2) \quad E_\epsilon(f) = \{x \in X : |f(x)| \geq \epsilon\}$$

has only finitely many elements.

Let $c_0(X, \mathbf{R})$ and $c_0(X, \mathbf{C})$ be the spaces of real and complex-valued functions on X that vanish at infinity, respectively. Remember that $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$ denote the spaces of real and complex-valued functions f on X such that the support of f has only finitely many elements, respectively, as in Section 1.6. In this case, f automatically vanishes at infinity on X , so that

$$(2.5.3) \quad c_{00}(X, \mathbf{R}) \subseteq c_0(X, \mathbf{R}), \quad c_{00}(X, \mathbf{C}) \subseteq c_0(X, \mathbf{C}).$$

In particular, if X has only finitely many elements, then every real or complex-valued function on X vanishes at infinity.

2.5.1 Sequences converging to 0

If f is a real or complex-valued function on the set \mathbf{Z}_+ of positive integers, then f vanishes at infinity on \mathbf{Z}_+ if and only if

$$(2.5.4) \quad \lim_{j \rightarrow \infty} f(j) = 0.$$

Let X be any nonempty set again, and let $\{x_j\}_{j=1}^\infty$ be an infinite sequence of distinct elements of X . Also let f be a real or complex-valued function on X , and suppose that the support of f is contained in the set of x_j 's, $j \in \mathbf{Z}_+$. Under these conditions, f vanishes at infinity on X if and only if

$$(2.5.5) \quad \lim_{j \rightarrow \infty} f(x_j) = 0.$$

Let f be any real or complex-valued function on X , and remember that the support of f is the set of $x \in X$ such that $f(x) \neq 0$. Equivalently,

$$(2.5.6) \quad \text{supp } f = \bigcup_{n=1}^{\infty} E_{1/n}(f),$$

where $E_{1/n}(f)$ is as in (2.5.2). If f vanishes at infinity on X , then it follows that the support of f has only finitely or countably many elements.

2.5.2 More on vanishing at infinity

If f is a real or complex-valued function on X that vanishes at infinity, then it is easy to see that f is bounded on X . In fact, we have that

$$(2.5.7) \quad \|f\|_\infty = \max_{x \in X} |f(x)|,$$

which is to say that the maximum on the right is attained. Of course, this is trivial when $f \equiv 0$ on X . Otherwise, if $f(x_0) \neq 0$ for some $x_0 \in X$, then there are only finitely many $x \in X$ such that $|f(x)| \geq |f(x_0)|$. In this case, it suffices to take the maximum of $|f(x)|$ over this finite set.

Thus

$$(2.5.8) \quad c_0(X, \mathbf{R}) \subseteq \ell^\infty(X, \mathbf{R}), \quad c_0(X, \mathbf{C}) \subseteq \ell^\infty(X, \mathbf{C}).$$

If g is another real or complex-valued function on X that vanishes at infinity, then one can check that $f + g$ vanishes at infinity on X as well. Similarly, if $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then tf vanishes at infinity on X . This means that $c_0(X, \mathbf{R})$ and $c_0(X, \mathbf{C})$ are linear subspaces of $\ell^\infty(X, \mathbf{R})$ and $\ell^\infty(X, \mathbf{C})$, respectively.

2.5.3 Closed sets in $\ell^\infty(X, \mathbf{R})$, $\ell^\infty(X, \mathbf{C})$

Let f be a bounded real or complex-valued function on X . Suppose that f is a limit point of $c_0(X, \mathbf{R})$ or $c_0(X, \mathbf{C})$ in $\ell^\infty(X, \mathbf{R})$ or $\ell^\infty(X, \mathbf{C})$, respectively, with

respect to the supremum metric. We would like to verify that f also vanishes at infinity on X in this case. Let $\epsilon > 0$ be given. By hypothesis, there is a real or complex-valued function g on X , as appropriate, such that g vanishes at infinity on X , and

$$(2.5.9) \quad \|f - g\|_\infty < \epsilon/2.$$

This implies that

$$(2.5.10) \quad |f(x)| \leq |g(x)| + |f(x) - g(x)| \leq |g(x)| + \|f - g\|_\infty < |g(x)| + \epsilon/2$$

for every $x \in X$. Of course, $|g(x)| < \epsilon/2$ for all but finitely many $x \in X$, because g vanishes at infinity on X . It follows that

$$(2.5.11) \quad |f(x)| < \epsilon/2 + \epsilon/2 = \epsilon$$

for all but finitely many $x \in X$, as desired.

This shows that $c_0(X, \mathbf{R})$ and $c_0(X, \mathbf{C})$ are closed sets in $\ell^\infty(X, \mathbf{R})$ and $\ell^\infty(X, \mathbf{C})$, respectively, with respect to the supremum metric. As a slightly different version of this, let $\{f_j\}_{j=1}^\infty$ be a sequence of real or complex-valued functions on X that vanish at infinity and converge uniformly to a real or complex-valued function f on X , as appropriate. Under these conditions, f vanishes at infinity on X too. This can be obtained from the previous statement, or using the same argument as in the preceding paragraph.

Let f be a real or complex-valued function on X that vanishes at infinity, and let $\epsilon > 0$ be given. Let f_ϵ be the real or complex-valued function, as appropriate, defined on X by

$$(2.5.12) \quad \begin{aligned} f_\epsilon(x) &= f(x) && \text{when } |f(x)| \geq \epsilon \\ &= 0 && \text{when } |f(x)| < \epsilon. \end{aligned}$$

Note that f_ϵ has finite support in X , because f vanishes at infinity on X . By construction,

$$(2.5.13) \quad \|f - f_\epsilon\|_\infty < \epsilon,$$

where the strict inequality uses the hypothesis that f vanishes at infinity on X again. This shows that $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$ are dense in $c_0(X, \mathbf{R})$ and $c_0(X, \mathbf{C})$, respectively, with respect to the supremum metric.

Chapter 3

Some more metric spaces

3.1 Continuous functions on closed intervals

Let a, b be real numbers with $a < b$, and let

$$(3.1.1) \quad [a, b] = \{x \in \mathbf{R} : a \leq x \leq b\}$$

be the usual closed interval in the real line from a to b . It is well known that $[a, b]$ is compact with respect to the standard metric on \mathbf{R} . As in Section 1.14, $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$ are the spaces of continuous real and complex-valued functions on $[a, b]$. This uses the restriction of the standard metric on \mathbf{R} to $[a, b]$, and the standard Euclidean metrics on \mathbf{R} and \mathbf{C} , as appropriate. As before, these spaces are linear subspaces of the real and complex vector spaces of all real and complex-valued functions on $[a, b]$, as appropriate.

3.1.1 Norms on $C([a, b], \mathbf{R})$, $C([a, b], \mathbf{C})$

As usual, a nonnegative real-valued function N on $C([a, b], \mathbf{R})$ or $C([a, b], \mathbf{C})$ is said to be a *norm* if it satisfies the following three conditions. First, $N(f) = 0$ if and only if $f = 0$. Second, if f is a continuous real or complex-valued function on $[a, b]$ and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then

$$(3.1.2) \quad N(tf) = |t|N(f).$$

Third, if f and g are continuous real or complex-valued functions on $[a, b]$, as appropriate, then

$$(3.1.3) \quad N(f + g) \leq N(f) + N(g).$$

In this case,

$$(3.1.4) \quad d_N(f, g) = N(f - g)$$

defines a metric on $C([a, b], \mathbf{R})$ or $C([a, b], \mathbf{C})$, as appropriate.

Remember that continuous real and complex-valued functions on $[a, b]$ are bounded, because $[a, b]$ is compact. The *supremum norm* is defined on each of $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$ by

$$(3.1.5) \quad \|f\|_{\infty} = \sup_{a \leq x \leq b} |f(x)|,$$

as in (1.13.4). Note that the supremum is attained in this situation, because of the extreme value theorem. The corresponding metric

$$(3.1.6) \quad d_{\infty}(f, g) = \|f - g\|_{\infty}$$

is the same as the supremum metric, as before.

3.1.2 The L^1 norm and metric on $C([a, b], \mathbf{R})$, $C([a, b], \mathbf{C})$

If f is a continuous real or complex-valued function on $[a, b]$, then put

$$(3.1.7) \quad \|f\|_1 = \int_a^b |f(x)| dx,$$

using the standard Riemann integral on the right side. This is the same as (1.15.4) in the case of the unit interval in \mathbf{R} . As before, one can check that (3.1.7) defines a norm on each of $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$. Let

$$(3.1.8) \quad d_1(f, g) = \|f - g\|_1 = \int_a^b |f(x) - g(x)| dx$$

be the metric associated to (3.1.7) on each of $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$.

If f is a continuous real or complex-valued function on $[a, b]$, then

$$(3.1.9) \quad \|f\|_1 = \int_a^b |f(x)| dx \leq (b - a) \|f\|_{\infty}.$$

Hence

$$(3.1.10) \quad d_1(f, g) = \|f - g\|_1 \leq (b - a) \|f - g\|_{\infty} = (b - a) d_{\infty}(f, g)$$

for all real or complex-valued continuous functions f and g on $[a, b]$. This implies that the identity mappings on $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$ are Lipschitz with respect to $d_{\infty}(f, g)$ on the domain and $d_1(f, g)$ on the range.

3.2 The square norm

Let a, b be real numbers with $a < b$ again. Put

$$(3.2.1) \quad \|f\|_2 = \left(\int_a^b |f(x)|^2 dx \right)^{1/2}$$

for each continuous real or complex-valued continuous function f on $[a, b]$, using the nonnegative square root on the right side, as usual. Note that this is the same as (1.15.6) in the case of the unit interval. It is easy to see that this satisfies the homogeneity and positivity requirements of a norm, as before. To get the triangle inequality, one can use an integral version of the Cauchy–Schwarz inequality, as follows.

3.2.1 The Cauchy–Schwarz inequality for integrals

Let f, g be continuous real or complex-valued functions on $[a, b]$. Observe that

$$(3.2.2) \quad |f(x)| |g(x)| \leq \frac{1}{2} |f(x)|^2 + \frac{1}{2} |g(x)|^2$$

for every $a \leq x \leq b$, as in (2.3.6). This implies that

$$(3.2.3) \quad \begin{aligned} \int_a^b |f(x)| |g(x)| dx &\leq \frac{1}{2} \int_a^b |f(x)|^2 dx + \frac{1}{2} \int_a^b |g(x)|^2 dx \\ &= \frac{1}{2} \|f\|_2^2 + \frac{1}{2} \|g\|_2^2. \end{aligned}$$

Using this, we can get that

$$(3.2.4) \quad \int_a^b |f(x)| |g(x)| dx \leq \|f\|_2 \|g\|_2,$$

which is the integral version of the Cauchy–Schwarz inequality mentioned in the preceding paragraph. More precisely, (3.2.4) follows from (3.2.3) when $\|f\|_2 = \|g\|_2 = 1$. If $\|f\|_2, \|g\|_2 > 0$, then one can reduce to the previous case, using scalar multiplication. Otherwise, (3.2.4) is trivial when $f = 0$ or $g = 0$ on $[a, b]$.

3.2.2 The triangle inequality for the L^2 norm

Clearly

$$(3.2.5) \quad \begin{aligned} \|f + g\|_2^2 &= \int_a^b |f(x) + g(x)|^2 dx \\ &\leq \int_a^b (|f(x)| + |g(x)|)^2 dx \\ &= \int_a^b |f(x)|^2 dx + 2 \int_a^b |f(x)| |g(x)| dx + \int_a^b |g(x)|^2 dx. \end{aligned}$$

It follows that

$$(3.2.6) \quad \|f + g\|_2^2 \leq \|f\|_2^2 + 2 \|f\|_2 \|g\|_2 + \|g\|_2^2 = (\|f\|_2 + \|g\|_2)^2,$$

using (3.2.4) in the first step. Thus

$$(3.2.7) \quad \|f + g\|_2 \leq \|f\|_2 + \|g\|_2,$$

so that (3.2.1) defines a norm on each of $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$. The associated metric is given by

$$(3.2.8) \quad d_2(f, g) = \|f - g\|_2 = \left(\int_a^b |f(x) - g(x)|^2 dx \right)^{1/2}.$$

3.2.3 More on the L^2 norm

If f is a continuous real or complex-valued function on $[a, b]$, then

$$(3.2.9) \quad \|f\|_2^2 = \int_a^b |f(x)|^2 dx \leq (b - a) \|f\|_\infty^2.$$

Equivalently,

$$(3.2.10) \quad \|f\|_2 \leq (b - a)^{1/2} \|f\|_\infty.$$

This implies that

$$(3.2.11) \quad d_2(f, g) \leq (b - a)^{1/2} d_\infty(f, g)$$

for all real or complex-valued continuous functions f and g on $[a, b]$. In particular, the identity mappings on $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$ are Lipschitz with respect to $d_\infty(f, g)$ on the domain and $d_2(f, g)$ on the range.

Observe that

$$(3.2.12) \quad \|f\|_1 = \int_a^b |f(x)| dx \leq (b - a)^{1/2} \|f\|_2$$

for every continuous real or complex-valued function f on $[a, b]$, by (3.2.4).

Hence

$$(3.2.13) \quad d_1(f, g) \leq (b - a)^{1/2} d_2(f, g)$$

for all continuous real or complex-valued functions f and g on $[a, b]$. It follows that the identity mappings on $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$ are Lipschitz with respect to $d_2(f, g)$ on the domain and $d_1(f, g)$ on the range.

3.3 Riemann–Stieltjes integrals

Let a, b be real numbers with $a < b$, and let α be a monotonically increasing real-valued function on $[a, b]$. If f is a continuous real-valued function on $[a, b]$, then the corresponding *Riemann–Stieltjes integral*

$$(3.3.1) \quad \int_a^b f d\alpha = \int_a^b f(x) d\alpha(x)$$

can be defined as a real number in a natural way. This reduces to the ordinary Riemann integral of f on $[a, b]$ when $\alpha(x) = x$ for every $x \in [a, b]$.

To be a bit more precise, let $\mathcal{P} = \{t_j\}_{j=0}^l$ be a *partition* of $[a, b]$. This is a finite sequence of real numbers with

$$(3.3.2) \quad a = t_0 < t_1 < \cdots < t_{l-1} < t_l = b,$$

as usual. In the definition of the Riemann integral, one considers certain sums associated to such a partition, involving the values of f on each interval $[t_{j-1}, t_j]$, and the length of the interval. In the definition of the Riemann–Stieltjes integral, one considers sums of a similar type, but using

$$(3.3.3) \quad \alpha(t_j) - \alpha(t_{j-1})$$

instead of the length of $[t_{j-1}, t_j]$. Note that (3.3.3) is greater than or equal to 0 for each j , and that

$$(3.3.4) \quad \sum_{j=1}^l (\alpha(t_j) - \alpha(t_{j-1})) = \alpha(b) - \alpha(a).$$

Riemann–Stieltjes integrals are discussed in many textbooks, as well as the article [119]. See [38] for a geometric interpretation of the Riemann–Stieltjes integral, and [206, 240] for some variants of the Riemann–Stieltjes integral.

3.3.1 Some basic properties of Riemann–Stieltjes integrals

The Riemann–Stieltjes integral has many of the same properties as the Riemann integral. In particular, (3.3.1) is greater than or equal to 0 when $f \geq 0$ on $[a, b]$. In this case, if $f(x_0) > 0$ for some $x_0 \in [a, b]$, and if α is not constant near x_0 , then (3.3.1) is positive. If α is strictly increasing on $[a, b]$, then this works for any $x_0 \in [a, b]$. If α is constant on $[a, b]$, then (3.3.1) is equal to 0 for any function f on $[a, b]$.

If α is continuously-differentiable on $[a, b]$, then it is well known that

$$(3.3.5) \quad \int_a^b f(x) d\alpha(x) = \int_a^b f(x) \alpha'(x) dx.$$

In fact, this works when α is differentiable at every point in $[a, b]$, and its derivative α' is Riemann integrable on $[a, b]$.

Although monotonically increasing real-valued functions on $[a, b]$ are not necessarily continuous, it is well known that their only possible discontinuities are jump discontinuities. It is also well known that such a function can have only finitely or countably many discontinuities in $[a, b]$. It is interesting to consider the Riemann–Stieltjes integral when α is a monotonically increasing step function on $[a, b]$, for instance.

3.3.2 Riemann–Stieltjes integrability

It is well known that the Riemann integral of a bounded real-valued function f on $[a, b]$ can be defined when f is Riemann integrable on $[a, b]$, in a suitable sense, which includes the case where f is continuous on $[a, b]$. Similarly, (3.3.1) can be defined when f is Riemann–Stieltjes integrable on $[a, b]$ with respect to α , in a suitable sense, which includes the case when f is continuous on $[a, b]$.

If f is a complex-valued function on $[a, b]$, then one can consider the corresponding integrability properties of the real and imaginary parts of f . If the real and imaginary parts of f are Riemann–Stieltjes integrable on $[a, b]$ with respect to α , then one can define the Riemann–Stieltjes integral of f as a complex number, whose real and imaginary parts are the corresponding Riemann–Stieltjes integrals of the real and imaginary parts of f . In particular, this works when f is continuous on $[a, b]$.

3.4 Riemann–Stieltjes integrals and seminorms

Let a, b be real numbers with $a < b$, and let N be a nonnegative real-valued function on $C([a, b], \mathbf{R})$ or $C([a, b], \mathbf{C})$. As in Subsection A.6.1, N is said to be a *seminorm* if

$$(3.4.1) \quad N(tf) = |t|N(f)$$

and

$$(3.4.2) \quad N(f + g) \leq N(f) + N(g)$$

for every $f, g \in C([a, b], \mathbf{R})$ or $C([a, b], \mathbf{C})$ and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate. In this case, $N(0) = 0$, by taking $t = 0$ in (3.4.1). If we also have that $N(f) > 0$ when $f \not\equiv 0$ on $[a, b]$, then N is a norm on $C([a, b], \mathbf{R})$ or $C([a, b], \mathbf{C})$, as appropriate.

3.4.1 Some integral seminorms

Let α be a monotonically increasing real-valued function on $[a, b]$. If f is a continuous real or complex-valued function on $[a, b]$, then

$$(3.4.3) \quad \|f\|_{1,\alpha} = \int_a^b |f(x)| d\alpha(x)$$

is defined as a nonnegative real number, as in the previous section. Of course, this is the same as (3.1.7) when $\alpha(x) = x$ on $[a, b]$. It is easy to see that

$$(3.4.4) \quad \|tf\|_{1,\alpha} = |t|\|f\|_{1,\alpha}$$

and

$$(3.4.5) \quad \|f + g\|_{1,\alpha} \leq \|f\|_{1,\alpha} + \|g\|_{1,\alpha}$$

for all continuous real or complex-valued functions f and g on $[a, b]$ and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate. Thus (3.4.3) defines a seminorm on each of $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$.

If α is strictly increasing on $[a, b]$, and if f is a continuous real or complex-valued function on $[a, b]$ such that $f(x_0) \neq 0$ for some $x_0 \in [a, b]$, then one can check that (3.4.3) is positive, as before. This implies that (3.4.3) defines a norm on each of $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$ in this case. It follows that

$$(3.4.6) \quad d_{1,\alpha}(f, g) = \|f - g\|_{1,\alpha}$$

defines a metric on each of $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$ under these conditions.

3.4.2 More on these integral seminorms

If f, g are continuous real or complex-valued function on $[a, b]$, then

$$(3.4.7) \quad \begin{aligned} \left| \int_a^b f d\alpha - \int_a^b g d\alpha \right| &= \left| \int_a^b (f - g) d\alpha \right| \\ &\leq \int_a^b |f - g| d\alpha = \|f - g\|_{1,\alpha}, \end{aligned}$$

using basic properties of the Riemann–Stieltjes integral in the first two steps. Suppose for the moment that α is strictly increasing on $[a, b]$. It follows from (3.4.7) that

$$(3.4.8) \quad f \mapsto \int_a^b f d\alpha$$

is Lipschitz as a mapping from $C([a, b], \mathbf{R})$ or $C([a, b], \mathbf{C})$ into \mathbf{R} or \mathbf{C} , as appropriate, with respect to (3.4.6) on the domain and the standard Euclidean metric on the range.

If f is a continuous real or complex-valued function on $[a, b]$, then

$$(3.4.9) \quad \|f\|_{1,\alpha} \leq (\alpha(b) - \alpha(a)) \|f\|_\infty,$$

where $\|f\|_\infty$ is the supremum norm of f on $[a, b]$, as in (3.1.5). Combining (3.4.7) and (3.4.9), we get that

$$(3.4.10) \quad \left| \int_a^b f d\alpha - \int_a^b g d\alpha \right| \leq (\alpha(b) - \alpha(a)) \|f - g\|_\infty,$$

for all continuous real or complex-valued functions f, g on $[a, b]$. This implies that (3.4.8) is Lipschitz as a mapping from $C([a, b], \mathbf{R})$ or $C([a, b], \mathbf{C})$ into \mathbf{R} or \mathbf{C} , as appropriate, with respect to the supremum metric on the domain and the standard Euclidean metric on the range.

3.4.3 Connections with uniform convergence

Let $\{f_j\}_{j=1}^\infty$ be a sequence of real or complex-valued functions on $[a, b]$ that converges uniformly to a real or complex-valued function f on $[a, b]$, as appropriate. If f_j is continuous on $[a, b]$ for each j , then f is continuous on $[a, b]$, and

$$(3.4.11) \quad \|f_j - f\|_{1,\alpha} \rightarrow 0 \text{ as } j \rightarrow \infty,$$

by (3.4.9). This means that $\{f_j\}_{j=1}^\infty$ converges to f with respect to (3.4.6) when α is strictly increasing on $[a, b]$. We also get that

$$(3.4.12) \quad \lim_{j \rightarrow \infty} \int_a^b f_j(x) d\alpha(x) = \int_a^b f(x) d\alpha(x),$$

by (3.4.10).

More precisely, if f_j is Riemann–Stieltjes integrable on $[a, b]$ with respect to α for each j , then it is well known that f is Riemann–Stieltjes integrable on $[a, b]$ with respect to α too. In this case, (3.4.12) holds for essentially the same reasons as before. Similarly, one can define (3.4.3) for Riemann–Stieltjes integrable functions on $[a, b]$ with respect to α , and (3.4.11) holds as well.

3.5 Square seminorms

Let a, b be real numbers with $a < b$, and let α be a monotonically increasing real-valued function on $[a, b]$. If f is a continuous real or complex-valued function on $[a, b]$, then

$$(3.5.1) \quad \|f\|_{2,\alpha} = \left(\int_a^b |f(x)|^2 d\alpha(x) \right)^{1/2}$$

is defined as a nonnegative real number. This reduces to (3.2.1) when $\alpha(x) = x$ on $[a, b]$, as before. Note that

$$(3.5.2) \quad \|t f\|_{2,\alpha} = |t| \|f\|_{2,\alpha}$$

for every $t \in \mathbf{R}$ or \mathbf{C} , as appropriate.

3.5.1 Another Cauchy–Schwarz inequality for integrals

Let g be another continuous real or complex-valued function on $[a, b]$. Observe that

$$(3.5.3) \quad \begin{aligned} \int_a^b |f(x)| |g(x)| d\alpha(x) &\leq \frac{1}{2} \int_a^b |f(x)|^2 d\alpha(x) + \frac{1}{2} \int_a^b |g(x)|^2 d\alpha(x) \\ &= \frac{1}{2} \|f\|_{2,\alpha}^2 + \frac{1}{2} \|g\|_{2,\alpha}^2, \end{aligned}$$

using (3.2.2) in the first step. In particular,

$$(3.5.4) \quad \int_a^b |f(x)| |g(x)| d\alpha(x) \leq 1$$

when $\|f\|_{2,\alpha}, \|g\|_{2,\alpha} \leq 1$.

Let f, g be arbitrary continuous real or complex-valued functions on $[a, b]$, and let r, t be positive real numbers such that

$$(3.5.5) \quad \|f\|_{2,\alpha} \leq r, \quad \|g\|_{2,\alpha} \leq t.$$

Thus $\|r^{-1} f\|_{2,\alpha}, \|t^{-1} g\|_{2,\alpha} \leq 1$, so that

$$(3.5.6) \quad \int_a^b |r^{-1} f(x)| |t^{-1} g(x)| d\alpha(x) \leq 1.$$

Equivalently, this means that

$$(3.5.7) \quad \int_a^b |f(x)| |g(x)| d\alpha(x) \leq r t.$$

One can use this to get that

$$(3.5.8) \quad \int_a^b |f(x)| |g(x)| d\alpha(x) \leq \|f\|_{2,\alpha} \|g\|_{2,\alpha},$$

which is the integral version of the Cauchy–Schwarz inequality for Riemann–Stieltjes integrals on $[a, b]$.

3.5.2 The triangle inequality for these square seminorms

As in Subsection 3.2.2, one can show that

$$(3.5.9) \quad \|f + g\|_{2,\alpha} \leq \|f\|_{2,\alpha} + \|g\|_{2,\alpha},$$

using (3.5.8). This implies that (3.5.1) defines a seminorm on each of $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$.

If α is strictly increasing on $[a, b]$, and if f is a continuous real or complex-valued function on $[a, b]$ that is not equal to 0 everywhere on $[a, b]$, then one can verify that (3.5.1) is positive, as usual. In this case, (3.5.1) defines a norm on each of $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$. Thus

$$(3.5.10) \quad d_{2,\alpha}(f, g) = \|f - g\|_{2,\alpha}$$

defines a metric on each of these spaces in this situation.

3.5.3 More on these square seminorms

Observe that

$$(3.5.11) \quad \|f\|_{2,\alpha} \leq (\alpha(b) - \alpha(a))^{1/2} \|f\|_{\infty}$$

for every continuous real or complex-valued function f on $[a, b]$. Of course, this implies that

$$(3.5.12) \quad \|f - g\|_{2,\alpha} \leq (\alpha(b) - \alpha(a))^{1/2} \|f - g\|_{\infty}$$

for all continuous real or complex-valued functions f, g on $[a, b]$.

Similarly,

$$(3.5.13) \quad \|f\|_{1,\alpha} \leq (\alpha(b) - \alpha(a))^{1/2} \|f\|_{2,\alpha},$$

by (3.5.8). It follows that

$$(3.5.14) \quad \|f - g\|_{1,\alpha} \leq (\alpha(b) - \alpha(a))^{1/2} \|f - g\|_{2,\alpha}$$

for every $f, g \in C([a, b], \mathbf{R})$ or $C([a, b], \mathbf{C})$.

3.6 Some more monotone functions

Let a, b be real numbers with $a < b$, and let α, β be monotonically increasing functions real-valued functions on $[a, b]$. Thus $\alpha + \beta$ is a monotonically increasing real-valued function on $[a, b]$ as well. If f is a continuous real-valued function on $[a, b]$, then the Riemann–Stieltjes integrals of f on $[a, b]$ with respect to $\alpha, \beta, \alpha + \beta$ may be defined in the usual way. It is well known that

$$(3.6.1) \quad \int_a^b f d(\alpha + \beta) = \int_a^b f d\alpha + \int_a^b f d\beta.$$

In particular, if $f \geq 0$ on $[a, b]$, then

$$(3.6.2) \quad \int_a^b f d\alpha, \int_a^b f d\beta \leq \int_a^b f d(\alpha + \beta).$$

Suppose now that f is a continuous real or complex-valued function on $[a, b]$, so that $\|f\|_{1,\alpha}$ and $\|f\|_{2,\alpha}$ may be defined as in the previous two sections. Of course, the analogous quantities

$$(3.6.3) \quad \|f\|_{1,\beta}, \|f\|_{2,\beta}, \|f\|_{1,\alpha+\beta}, \|f\|_{2,\alpha+\beta}$$

associated to β and $\alpha + \beta$ may be defined in the same way. Observe that

$$(3.6.4) \quad \|f\|_{1,\alpha+\beta} = \|f\|_{1,\alpha} + \|f\|_{1,\beta},$$

because of (3.6.1). Similarly,

$$(3.6.5) \quad \|f\|_{2,\alpha+\beta}^2 = \|f\|_{2,\alpha}^2 + \|f\|_{2,\beta}^2.$$

It follows that

$$(3.6.6) \quad \|f\|_{1,\alpha}, \|f\|_{1,\beta} \leq \|f\|_{1,\alpha+\beta}$$

and

$$(3.6.7) \quad \|f\|_{2,\alpha}, \|f\|_{2,\beta} \leq \|f\|_{2,\alpha+\beta}$$

3.6.1 Some homogeneity properties

Let t be a nonnegative real number, so that $t\alpha$ is a monotonically increasing real-valued function on $[a, b]$ too. If f is a continuous real-valued function on $[a, b]$, then the Riemann–Stieltjes integral of f on $[a, b]$ with respect to $t\alpha$ may be defined in the usual way. It is well known that

$$(3.6.8) \quad \int_a^b f d(t\alpha) = t \int_a^b f d\alpha.$$

If f is a continuous real or complex-valued function on $[a, b]$, then

$$(3.6.9) \quad \|f\|_{1,t\alpha} \text{ and } \|f\|_{2,t\alpha}$$

may be defined as in the two previous sections again. It is easy to see that

$$(3.6.10) \quad \|f\|_{1,t\alpha} = t \|f\|_{1,\alpha}$$

and

$$(3.6.11) \quad \|f\|_{2,t\alpha} = \sqrt{t} \|f\|_{2,\alpha},$$

using (3.6.8).

3.7 Continuous functions on \mathbf{R}

In this section, we take the real line to be equipped with the standard Euclidean metric, as usual. Let f be a real or complex-valued function on \mathbf{R} , so that the *support* of f may be defined as the closure in \mathbf{R} of the set on which $f \neq 0$, as in Section A.11. If f has compact support in \mathbf{R} , then there are real numbers a, b with $a < b$ and

$$(3.7.1) \quad \text{supp } f \subseteq [a, b],$$

because compact subsets of \mathbf{R} are bounded. Conversely, if (3.7.1) holds for some $a, b \in \mathbf{R}$ with $a \leq b$, then $\text{supp } f$ is compact, because $[a, b]$ is a compact subset of \mathbf{R} , and because the support of f is a closed set by construction. The spaces $C_{com}(\mathbf{R}, \mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$ of continuous real and complex-valued functions on \mathbf{R} with compact support are linear subspaces of the real and complex vector spaces of all continuous real and complex-valued functions on \mathbf{R} , respectively, as in Subsection A.11.3.

3.7.1 Norms on $C_{com}(\mathbf{R}, \mathbf{R})$, $C_{com}(\mathbf{R}, \mathbf{C})$

A nonnegative real-valued function N on $C_{com}(\mathbf{R}, \mathbf{R})$ or $C_{com}(\mathbf{R}, \mathbf{C})$ is said to be a *norm* if it satisfies the following three conditions, as usual. First, $N(f) = 0$ if and only if $f = 0$. Second, if f is a continuous real or complex-valued function on \mathbf{R} with compact support and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then

$$(3.7.2) \quad N(tf) = |t| N(f).$$

Third, if f and g are continuous real or complex-valued functions on \mathbf{R} with compact support, as appropriate, then

$$(3.7.3) \quad N(f + g) \leq N(f) + N(g).$$

This implies that

$$(3.7.4) \quad d_N(f, g) = N(f - g)$$

defines a metric on $C_{com}(\mathbf{R}, \mathbf{R})$ or $C_{com}(\mathbf{R}, \mathbf{C})$, as appropriate.

Continuous real and complex-valued functions on \mathbf{R} with compact support are bounded, as mentioned in Subsection A.11.3. The *supremum norm* is defined on each of $C_{com}(\mathbf{R}, \mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$ by

$$(3.7.5) \quad \|f\|_{\infty} = \sup_{x \in \mathbf{R}} |f(x)|,$$

as in (1.13.4). Equivalently,

$$(3.7.6) \quad \|f\|_\infty = \sup\{|f(x)| : x \in \text{supp } f\}$$

when $\text{supp } f \neq \emptyset$, in which case the supremum is attained, by the extreme value theorem. If (3.7.1) holds for some $a, b \in \mathbf{R}$ with $a \leq b$, then

$$(3.7.7) \quad \|f\|_\infty = \sup_{a \leq x \leq b} |f(x)|.$$

As before, the corresponding metric

$$(3.7.8) \quad d_\infty(f, g) = \|f - g\|_\infty$$

is the same as the supremum metric.

3.7.2 Some integrals on \mathbf{R}

Let f be a continuous real or complex-valued function on \mathbf{R} with compact support, and let a, b be real numbers such that $a \leq b$ and (3.7.1) holds. Under these conditions, one can define the integral

$$(3.7.9) \quad \int_{-\infty}^{\infty} f(x) dx$$

to be the Riemann integral

$$(3.7.10) \quad \int_a^b f(x) dx.$$

More precisely, one can check that (3.7.10) does not depend on the particular choices of a and b , as long as (3.7.1) holds.

If g is another continuous real or complex-valued function on \mathbf{R} , as appropriate, with compact support, then

$$(3.7.11) \quad \int_{-\infty}^{\infty} (f(x) + g(x)) dx = \int_{-\infty}^{\infty} f(x) dx + \int_{-\infty}^{\infty} g(x) dx.$$

To see this, one can choose $a, b \in \mathbf{R}$ such that $a < b$ and the supports of both f and g are contained in $[a, b]$, so that the support of $f + g$ is contained in $[a, b]$ too. In this case, (3.7.11) reduces to the well-known fact that

$$(3.7.12) \quad \int_a^b (f(x) + g(x)) dx = \int_a^b f(x) dx + \int_a^b g(x) dx.$$

Of course, other basic properties of (3.7.9) follow from the corresponding statements for (3.7.10).

3.7.3 Some integral norms on $C_{com}(\mathbf{R}, \mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$

If f is a continuous real or complex-valued function on \mathbf{R} with compact support, then put

$$(3.7.13) \quad \|f\|_1 = \int_{-\infty}^{\infty} |f(x)| dx,$$

where the integral on the right is defined as in the preceding subsection. One can check that this defines a norm on each of $C_{com}(\mathbf{R}, \mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$, in essentially the same way as for continuous functions on an interval. This leads to a metric

$$(3.7.14) \quad d_1(f, g) = \|f - g\|_1 = \int_{-\infty}^{\infty} |f(x) - g(x)| dx$$

on each of $C_{com}(\mathbf{R}, \mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$, as in (3.7.4).

Similarly, put

$$(3.7.15) \quad \|f\|_2 = \left(\int_{-\infty}^{\infty} |f(x)|^2 dx \right)^{1/2}$$

for every continuous real or complex-valued function f on \mathbf{R} with compact support, using the nonnegative square root on the right side. One can verify that this defines a norm on each of $C_{com}(\mathbf{R}, \mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$, using the analogous statements for continuous functions on an interval. Thus

$$(3.7.16) \quad d_2(f, g) = \left(\int_{-\infty}^{\infty} |f(x) - g(x)|^2 dx \right)^{1/2}$$

defines a metric on each of $C_{com}(\mathbf{R}, \mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$.

Let f be any continuous real or complex-valued function on \mathbf{R} with compact support. Observe that

$$(3.7.17) \quad \|f\|_2^2 = \int_{-\infty}^{\infty} |f(x)|^2 dx \leq \|f\|_{\infty} \int_{-\infty}^{\infty} |f(x)| dx = \|f\|_1 \|f\|_{\infty}.$$

3.8 Riemann–Stieltjes integrals on \mathbf{R}

Let f be a continuous real-valued function on the real line with compact support, and let a, b be real numbers such that $a \leq b$ and (3.7.1) holds. Also let α be a monotonically increasing real-valued function on \mathbf{R} . The integral

$$(3.8.1) \quad \int_{-\infty}^{\infty} f d\alpha = \int_{-\infty}^{\infty} f(x) d\alpha(x)$$

may be defined as the Riemann–Stieltjes integral

$$(3.8.2) \quad \int_a^b f d\alpha.$$

This is the same as (3.7.9) when $\alpha(x) = x$ for all $x \in \mathbf{R}$, and this does not depend on the particular choices of a and b as long as (3.7.1) holds, as before.

If f is a continuous complex-valued function on \mathbf{R} , then (3.8.1) may be defined analogously using the extension of the Riemann–Stieltjes integral to complex-valued functions, as in Subsection 3.3.2, or equivalently by integrating the real and imaginary parts individually, as before.

Many basic properties of Riemann–Stieltjes integrals on closed intervals extend easily to this version, as in the previous section. In particular, (3.8.1) is linear in f , as before.

3.8.1 Some more integral seminorms

As in Section 3.4 and Subsection A.6.1, a nonnegative real-valued function N on $C_{com}(\mathbf{R}, \mathbf{R})$ or $C_{com}(\mathbf{R}, \mathbf{C})$ is said to be a *seminorm* if it satisfies (3.7.2) and (3.7.3). In particular, we can take $t = 0$ in (3.7.2), to get that $N(0) = 0$, as before.

It is easy to see that

$$(3.8.3) \quad \|f\|_{1,\alpha} = \int_{-\infty}^{\infty} |f(x)| d\alpha(x)$$

defines a seminorm on each of $C_{com}(\mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$, as in Subsection 3.4.1. If α is strictly increasing on \mathbf{R} , then this is a norm on $C_{com}(\mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$, as before. In this case,

$$(3.8.4) \quad d_{1,\alpha}(f, g) = \|f - g\|_{1,\alpha}$$

defines a metric on each of $C_{com}(\mathbf{R}, \mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$, as in (3.7.4).

Similarly,

$$(3.8.5) \quad \|f\|_{2,\alpha} = \left(\int_{-\infty}^{\infty} |f(x)|^2 d\alpha(x) \right)^{1/2}$$

defines a seminorm on each of $C_{com}(\mathbf{R}, \mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$, as in Section 3.5. If α is strictly increasing on \mathbf{R} , then this is a norm on $C_{com}(\mathbf{R}, \mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$. This means that

$$(3.8.6) \quad d_{2,\alpha}(f, g) = \|f - g\|_{2,\alpha}$$

is a metric on each of $C_{com}(\mathbf{R}, \mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$ under these conditions.

If f is a continuous real or complex-valued function on the real line with compact support, then we also have that

$$(3.8.7) \quad \|f\|_{2,\alpha}^2 = \int_{-\infty}^{\infty} |f(x)|^2 d\alpha(x) \leq \|f\|_{\infty} \int_{-\infty}^{\infty} |f(x)| d\alpha(x) = \|f\|_{1,\alpha} \|f\|_{\infty},$$

as before.

3.8.2 Some additional monotone functions

Let β be another monotonically increasing real-valued function on the real line, so that $\alpha + \beta$ is monotonically increasing on \mathbf{R} as well. If f is a continuous real or complex-valued function on \mathbf{R} with compact support, then

$$(3.8.8) \quad \int_{-\infty}^{\infty} f d(\alpha + \beta) = \int_{-\infty}^{\infty} f d\alpha + \int_{-\infty}^{\infty} f d\beta,$$

as in Section 3.6. In particular, if f is real-valued and nonnegative on \mathbf{R} , then

$$(3.8.9) \quad \int_{-\infty}^{\infty} f d\alpha, \int_{-\infty}^{\infty} f d\beta \leq \int_{-\infty}^{\infty} f d(\alpha + \beta),$$

as before.

We also have that

$$(3.8.10) \quad \|f\|_{1,\alpha+\beta} = \|f\|_{1,\alpha} + \|f\|_{1,\beta}$$

and

$$(3.8.11) \quad \|f\|_{2,\alpha+\beta}^2 = \|f\|_{2,\alpha}^2 + \|f\|_{2,\beta}^2,$$

as in Section 3.6. This implies that

$$(3.8.12) \quad \|f\|_{1,\alpha}, \|f\|_{1,\beta} \leq \|f\|_{1,\alpha+\beta}$$

and

$$(3.8.13) \quad \|f\|_{2,\alpha}, \|f\|_{2,\beta} \leq \|f\|_{2,\alpha+\beta},$$

as before.

If t is a nonnegative real number, then $t\alpha$ is another monotonically increasing real-valued function on \mathbf{R} , and

$$(3.8.14) \quad \int_{-\infty}^{\infty} f d(t\alpha) = t \int_{-\infty}^{\infty} f d\alpha,$$

as in Subsection 3.6.1. This implies that

$$(3.8.15) \quad \|f\|_{1,t\alpha} = t \|f\|_{1,\alpha}$$

and

$$(3.8.16) \quad \|f\|_{2,t\alpha} = \sqrt{t} \|f\|_{2,\alpha},$$

as before.

3.8.3 Another version of the Cauchy–Schwarz inequality

Let g be another continuous real or complex-valued function on the real line with compact support. Another version of the Cauchy–Schwarz inequality states that

$$(3.8.17) \quad \|fg\|_{1,\alpha} \leq \|f\|_{2,\alpha} \|g\|_{2,\alpha}.$$

This follows from the analogous statement for integrals over closed intervals in \mathbf{R} , as usual.

3.8.4 The case of bounded α

Let us suppose from now on in this section that α is bounded on \mathbf{R} . It is convenient to put

$$(3.8.18) \quad \alpha(\infty) = \sup\{\alpha(x) : x \in \mathbf{R}\}.$$

and

$$(3.8.19) \quad \alpha(-\infty) = \inf\{\alpha(x) : x \in \mathbf{R}\}$$

These may be considered as the limits of $\alpha(x)$ as $x \rightarrow \infty$ and $x \rightarrow -\infty$, respectively, because α is monotonically increasing on \mathbf{R} , by hypothesis. Note that

$$(3.8.20) \quad \alpha(-\infty) \leq \alpha(\infty),$$

with equality exactly when α is constant on \mathbf{R} .

If f is a continuous real or complex-valued function on \mathbf{R} with compact support again, then it is easy to see that

$$(3.8.21) \quad \|f\|_{1,\alpha} \leq (\alpha(\infty) - \alpha(-\infty)) \|f\|_{\infty},$$

because of the analogous statement for functions on closed intervals in \mathbf{R} . Similarly,

$$(3.8.22) \quad \|f\|_{2,\infty} \leq (\alpha(\infty) - \alpha(-\infty))^{1/2} \|f\|_{\infty}.$$

We also have that

$$(3.8.23) \quad \|f\|_{1,\alpha} \leq (\alpha(\infty) - \alpha(-\infty))^{1/2} \|f\|_{2,\alpha}.$$

3.9 Dense sets and Weierstrass' theorem

Let $(M, d(x, y))$ be a metric space, and let E be a subset of M . Remember that E is said to be *dense* in M if every element of M is an element of E , or a limit point of E , or both. This is the same as saying that the closure \overline{E} of E in M is equal to M .

Equivalently, it is easy to see that E is dense in M if and only if for every $x \in M$ and $\epsilon > 0$ there is a $y \in E$ such that

$$(3.9.1) \quad d(x, y) < \epsilon.$$

Alternatively, one can verify that E is dense in M if and only if for every $x \in M$ there is a sequence $\{x_j\}_{j=1}^{\infty}$ of elements of E that converges to x in M . More precisely, if E satisfies the condition in the preceding paragraph, then for each positive integer j one can choose a point $x_j \in E$ such that

$$(3.9.2) \quad d(x, x_j) < 1/j.$$

One can check that this implies that $\{x_j\}_{j=1}^{\infty}$ converges to x in M . Conversely, the existence of such a sequence clearly implies the condition in the preceding paragraph.

3.9.1 Density in $C_b(X, Y)$

Let (X, d_X) and (Y, d_Y) be nonempty metric spaces, and let $\theta(\cdot, \cdot)$ be the corresponding supremum metric on the space $C_b(X, Y)$ of bounded continuous mappings from X into Y , as in Subsection 1.11 and Section 1.14. A subset E of $C_b(X, Y)$ is dense in $C_b(X, Y)$ with respect to $\theta(\cdot, \cdot)$ if and only if for every $f \in C_b(X, Y)$ and $\epsilon > 0$ there is a $g \in E$ such that

$$(3.9.3) \quad \theta(f, g) < \epsilon.$$

This follows from the analogous remark before, with $M = C_b(X, Y)$.

Note that (3.9.3) implies that

$$(3.9.4) \quad d_Y(f(x), g(x)) < \epsilon \text{ for every } x \in X.$$

If (3.9.4) holds, then

$$(3.9.5) \quad \theta(f, g) \leq \epsilon,$$

by the definition of the supremum metric. Using this, we get that E is dense in $C_b(X, Y)$ with respect to $\theta(\cdot, \cdot)$ if and only if for every $f \in C_b(X, Y)$ and $\epsilon > 0$ there is a $g \in E$ such that (3.9.4) holds.

We also have that E is dense in $C_b(X, Y)$ with respect to $\theta(\cdot, \cdot)$ if and only if for every $f \in C_b(X, Y)$ there is a sequence $\{f_j\}_{j=1}^\infty$ of elements of E that converges to f uniformly on X . This uses the equivalence of uniform convergence and convergence with respect to $\theta(\cdot, \cdot)$ for sequences of bounded mappings from X into Y , as in Subsection 1.11.3.

3.9.2 Weierstrass' theorem

Let a, b be real numbers with $a < b$, and let f be a continuous complex-valued function on $[a, b]$. Of course, this uses the restriction of the standard Euclidean metric on \mathbf{R} to $[a, b]$, and the standard Euclidean metric on \mathbf{C} .

A famous theorem of Weierstrass states that there is a sequence of polynomials on \mathbf{R} with complex coefficients that converges to f uniformly on $[a, b]$. If f is a continuous real-valued function on $[a, b]$, then one can use polynomials on \mathbf{R} with real coefficients.

This means that the set of functions on $[a, b]$ corresponding to polynomials with complex coefficients is dense in the space $C([a, b], \mathbf{C})$ of continuous complex-valued functions on $[a, b]$, with respect to the supremum metric. Similarly, the set of functions on $[a, b]$ corresponding to polynomials with real coefficients is dense in the space $C([a, b], \mathbf{R})$ of continuous real-valued functions on $[a, b]$, with respect to the supremum metric.

Suppose now that (X, d_X) is a nonempty compact metric space. The Stone-Weierstrass theorem gives a criterion for a set of continuous real or complex-valued functions on X to be dense in $C(X, \mathbf{R})$ or $C(X, \mathbf{C})$, as appropriate, with respect to the supremum metric. This will be discussed in Section 3.12, after some preliminaries in the next two sections.

3.10 Functions on sets

Let X be a nonempty set, and let

$$(3.10.1) \quad c(X, \mathbf{R}), \quad c(X, \mathbf{C})$$

be the spaces of all real and complex-valued functions on X , respectively. Of course, if f and g are real or complex-valued functions on X , then $f + g$ is a real or complex-valued function on X , as appropriate. Similarly, if f is a real or complex-valued function on X , and t is a real or complex number, as appropriate, then tf is a real or complex-valued function on X , as appropriate. These are basic (classes of) examples of vector spaces over the real or complex numbers.

3.10.1 Linear subspaces of $c(X, \mathbf{R})$, $c(X, \mathbf{C})$

A subset W of $c(X, \mathbf{R})$ or $c(X, \mathbf{C})$ is said to be a *linear subspace* of $c(X, \mathbf{R})$ or $c(X, \mathbf{C})$, as appropriate, if it satisfies the following three conditions. First, $0 \in W$. Second, if $f, g \in W$, then

$$(3.10.2) \quad f + g \in W.$$

Third, if $f \in W$ and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then

$$(3.10.3) \quad tf \in W.$$

Note that the third condition implies the first condition when W is nonempty, by taking $t = 0$.

If f and g are real or complex-valued functions on X , then their product $f(x)g(x)$ defines a real or complex-valued function on X , as appropriate, too. In fact, $c(X, \mathbf{R})$, $c(X, \mathbf{C})$ are basic (classes of) examples of *commutative rings*, although we shall not discuss this in detail here. More precisely, these are *commutative algebras* over the real and complex numbers, respectively.

3.10.2 Subalgebras of $c(X, \mathbf{R})$, $c(X, \mathbf{C})$

Let \mathcal{A} be a linear subspace of $c(X, \mathbf{R})$ or $c(X, \mathbf{C})$. Suppose that for every $f, g \in \mathcal{A}$, we have that

$$(3.10.4) \quad fg \in \mathcal{A}.$$

Under these conditions, \mathcal{A} is said to be a *subalgebra* of $c(X, \mathbf{R})$ or $c(X, \mathbf{C})$, as appropriate. This basically corresponds to the first part of Definition 7.28 on p161 of [210].

If X is equipped with a metric, then the corresponding spaces $C(X, \mathbf{R})$ and $C(X, \mathbf{C})$ of continuous real and complex-valued functions on X are subalgebras of $c(X, \mathbf{R})$ and $c(X, \mathbf{C})$, respectively. If X is equipped with the discrete metric, then any mapping from X into another metric space is continuous. In particular, this means that

$$(3.10.5) \quad C(X, \mathbf{R}) = c(X, \mathbf{R}) \text{ and } C(X, \mathbf{C}) = c(X, \mathbf{C})$$

in this case.

If W is a linear subspace of $c(X, \mathbf{R})$ or $c(X, \mathbf{C})$, then a subset E of W may be called a linear subspace of W if E is a linear subspace of $c(X, \mathbf{R})$ or $c(X, \mathbf{C})$, as appropriate. Similarly, if \mathcal{A} is a subalgebra of $c(X, \mathbf{R})$ or $c(X, \mathbf{C})$, then a subset \mathcal{E} of \mathcal{A} may be called a subalgebra of \mathcal{A} if \mathcal{E} is a subalgebra of $c(X, \mathbf{R})$ or $c(X, \mathbf{C})$, as appropriate. In particular, if X is equipped with a metric, then the notions of linear subspaces and subalgebras of $C(X, \mathbf{R})$ and $C(X, \mathbf{C})$ may be defined in this way.

3.10.3 Invariance under complex conjugation

Suppose now that W is a subset of $c(X, \mathbf{C})$. We shall sometimes be concerned with situations where

(3.10.6) W contains the complex conjugate of each of its elements.

If W is a linear subspace of $c(X, \mathbf{C})$, then this is equivalent to the condition that

(3.10.7) W contain the real and imaginary parts of all of its elements.

This uses the fact that the real and imaginary parts of a complex number a are equal to $(a + \bar{a})/2$ and $(a - \bar{a})/(2i)$, respectively, where \bar{a} is the complex conjugate of a , as usual.

3.11 More on collections of functions

Let X be a nonempty set again, and let W be a subset of $c(X, \mathbf{R})$ or $c(X, \mathbf{C})$. We say that W is *nowhere vanishing* on X if for every $x \in X$ there is an $f \in W$ such that

(3.11.1) $f(x) \neq 0$.

In particular, this holds when W contains a nonzero constant function on X .

We say that W *separates points* in X if for every $x, w \in X$ with $x \neq w$ there is an $f \in W$ such that

(3.11.2) $f(x) \neq f(w)$.

Of course, at least one of $f(x)$ and $f(w)$ is not equal to 0 in this case.

3.11.1 Some continuous real-valued functions

Suppose for the moment that X is equipped with a metric $d(\cdot, \cdot)$. If $p \in X$, then one can check that

(3.11.3) $f_p(x) = d(x, p)$

is continuous as a real-valued function on X . More precisely, f_p is Lipschitz with constant $C = 1$ on X , with respect to the standard Euclidean metric on \mathbf{R} . In particular, f_p is uniformly continuous on X .

It follows that $C(X, \mathbf{R})$ separates points in X .

3.11.2 More on subalgebras of $c(X, \mathbf{R})$, $c(X, \mathbf{C})$

Let \mathcal{A} be a subalgebra of $c(X, \mathbf{R})$, and suppose that \mathcal{A} is nowhere vanishing on X , and that \mathcal{A} separates points in X . If $x, w \in X$ and $x \neq w$, then there is an $f \in \mathcal{A}$ such that $f(x)$ and $f(w)$ are any two given real numbers. This corresponds to Theorem 7.31 on p162 of [210].

Similarly, if \mathcal{A} is a subalgebra of $c(X, \mathbf{C})$ that is nowhere vanishing on X and that separates points in X , and if $x, w \in X$ and $x \neq w$, then there is an $f \in \mathcal{A}$ such that $f(x)$ and $f(w)$ are any two given complex numbers, as in [210].

More precisely, in both cases one can find elements of \mathcal{A} equal to 1 at x and to 0 at w , and equal to 0 at x and to 1 at w . One can use linear combinations of such functions to get elements of \mathcal{A} with arbitrary values at x and w .

If E is a nonempty finite subset of X , then for each $y \in E$ one can find an element of \mathcal{A} that is equal to 1 at y , and to 0 at every other element of E . Indeed, if $x \in E$ and $x \neq y$, then there is an element of \mathcal{A} that is equal to 1 at y and 0 at x , as before. The product of such functions is an element of \mathcal{A} that is equal to 1 at y and to 0 at every other element of E . One can use linear combinations of these functions to get elements of \mathcal{A} with arbitrary values at points in E .

3.11.3 Subalgebras and polynomials

Let p be a polynomial with real coefficients. If f is any real-valued function on X , then

$$(3.11.4) \quad (p \circ f)(x) = p(f(x))$$

defines another real-valued function on X .

Let \mathcal{A} be a subalgebra of $c(X, \mathbf{R})$. Suppose that $p(0) = 0$, which is the same as saying that the constant term in p is equal to 0. If $f \in \mathcal{A}$, then it is easy to see that

$$(3.11.5) \quad p \circ f \in \mathcal{A}.$$

More precisely, $f^l \in \mathcal{A}$ for each positive integer l , and $p \circ f$ is the same as a linear combination of positive integer powers of f . This also works without asking that $p(0) = 0$ if \mathcal{A} contains the constant functions on X .

If p is a polynomial with complex coefficients, then one can think of p as defining a complex-valued function on the complex plane. In this case, if f is a complex-valued function on X , then $p \circ f$ defines a complex-valued function on X too, as in (3.11.4).

If \mathcal{A} is a subalgebra of $c(X, \mathbf{C})$ and $p(0) = 0$, then (3.11.5) holds for every $f \in \mathcal{A}$, for the same reasons as before. This also works without asking that $p(0) = 0$ when \mathcal{A} contains the constant functions on X , as before.

3.12 The Stone–Weierstrass theorem

Let $(X, d(x, y))$ be a metric space, and suppose that X is nonempty and compact. Also let \mathcal{A} be a subalgebra of the algebra $C(X, \mathbf{R})$ of continuous real-

valued functions on X , and suppose that

$$(3.12.1) \quad \mathcal{A} \text{ is nowhere vanishing on } X,$$

and that

$$(3.12.2) \quad \mathcal{A} \text{ separates points in } X.$$

Under these conditions, the *Stone–Weierstrass theorem* states that

$$(3.12.3) \quad \mathcal{A} \text{ is dense in } C(X, \mathbf{R}),$$

with respect to the supremum metric. It is easy to see that (3.12.1) and (3.12.2) are necessary for (3.12.3) to hold, because constant functions are continuous on X , and $C(X, \mathbf{R})$ separates points in X .

Some aspects of the proof of the Stone–Weierstrass theorem will be discussed in the next two sections.

3.12.1 Some remarks about Weierstrass' theorem

Let a, b be real numbers with $a < b$. Observe that the collection of functions on $[a, b]$ that can be expressed as a polynomial with real coefficients is a subalgebra of the algebra of continuous real-valued functions on $[a, b]$. It is easy to see that this subalgebra is nowhere vanishing on $[a, b]$, and that it separates points in $[a, b]$.

Thus Weierstrass' original approximation theorem for approximating real-valued continuous functions on $[a, b]$ by polynomials with real coefficients uniformly on $[a, b]$ can be obtained from the Stone–Weierstrass theorem. However, Weierstrass' original theorem, or at least an interesting case of it, is typically used to show the Stone–Weierstrass theorem.

3.12.2 The complex version of the Stone–Weierstrass theorem

Let $(X, d(x, y))$ be any nonempty compact metric space again, and now let \mathcal{A} be a subalgebra of the algebra $C(X, \mathbf{C})$ of continuous complex-valued functions on X . Suppose that \mathcal{A} is nowhere vanishing on X , that \mathcal{A} separates points in X , and that

$$(3.12.4) \quad \mathcal{A} \text{ contains the complex conjugate of each of its elements.}$$

In this case, another version of the Stone–Weierstrass theorem states that

$$(3.12.5) \quad \mathcal{A} \text{ is dense in } C(X, \mathbf{C}),$$

with respect to the supremum metric.

This version of the Stone–Weierstrass theorem can be obtained from the previous one, as follows. Note that

$$(3.12.6) \quad \mathcal{A} \text{ contains the real and imaginary parts of all of its elements,}$$

as in Subsection 3.10.3. One can check that

$$(3.12.7) \quad \text{the collection of real-valued functions in } \mathcal{A}$$

is a subalgebra of $C(X, \mathbf{R})$ that satisfies (3.12.1) and (3.12.2). Thus the previous version of the theorem implies that every continuous real-valued function on X can be approximated by elements of (3.12.7), uniformly on X . Using this, it is easy to see that every continuous complex-valued function on X can be approximated by elements of \mathcal{A} , uniformly on X .

It is well known that this version of the Stone–Weierstrass theorem does not always work without the hypothesis (3.12.4). One can get a counterexample by taking X to be the closed unit disk in the complex plane, and \mathcal{A} to be the algebra of continuous complex-valued functions on the closed unit disk defined by polynomials in a complex variable with coefficients in \mathbf{C} . It can be shown that (3.12.5) does not hold in this case, using complex analysis.

Alternatively, one can take X to be the unit circle in the complex plane, and \mathcal{A} to be the algebra of continuous complex-valued functions on the circle defined by polynomials in a complex variable with complex coefficients. One can show that (3.12.5) does not hold in this case without using complex analysis, as in Exercise 21 on p169 of [210].

3.13 Algebras of bounded functions

Let X be a nonempty set, and remember that $\ell^\infty(X, \mathbf{R})$ and $\ell^\infty(X, \mathbf{C})$ denote the spaces of bounded real and complex-valued functions on X , respectively, as in Section 1.13. If f, g are bounded real or complex-valued functions on X , then it is easy to see that their product fg is bounded on X as well. More precisely, if $x \in X$, then

$$(3.13.1) \quad |f(x)g(x)| = |f(x)| |g(x)| \leq \|f\|_\infty \|g\|_\infty.$$

This implies that the supremum norm of fg on X satisfies

$$(3.13.2) \quad \|fg\|_\infty \leq \|f\|_\infty \|g\|_\infty.$$

Thus $\ell^\infty(X, \mathbf{R})$ and $\ell^\infty(X, \mathbf{C})$ are subalgebras of the algebras $c(X, \mathbf{R})$ and $c(X, \mathbf{C})$ of all real and complex-valued functions on X , respectively, as in Subsection 3.10.2.

3.13.1 Some remarks about closures

Let W be a collection of real or complex-valued functions on X . If W contains the limits of all sequences of its elements that converge uniformly on X , then one may say that W is *uniformly closed*, as on p161 of [210].

Consider the collection of all real or complex-valued functions on X that may be obtained as limits of sequences of elements of W that converge uniformly on X . This may be described as the *uniform closure* of W , as on p161 of [210].

If W is a subset of $\ell^\infty(X, \mathbf{R})$ or $\ell^\infty(X, \mathbf{C})$, then let \overline{W} be the closure of W in $\ell^\infty(X, \mathbf{R})$ or $\ell^\infty(X, \mathbf{C})$, as appropriate, with respect to the supremum metric. One can verify that this is the same as the uniform closure of W , as in the preceding paragraph.

If W is a linear subspace of $c(X, \mathbf{R})$ or $c(X, \mathbf{C})$, then one can check that the uniform closure of W is a linear subspace of $c(X, \mathbf{R})$ or $c(X, \mathbf{C})$, as appropriate. In fact, if $\{f_j\}_{j=1}^\infty$ and $\{g_j\}_{j=1}^\infty$ are sequences of real or complex-valued functions on X that converge uniformly to real or complex-valued functions f and g on X , respectively, then it is easy to see that $\{f_j + g_j\}_{j=1}^\infty$ converges to $f + g$ uniformly on X . Similarly, if $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then $\{t f_j\}_{j=1}^\infty$ converges to $t f$ uniformly on X . In particular, if W is a linear subspace of $\ell^\infty(X, \mathbf{R})$ or $\ell^\infty(X, \mathbf{C})$, then

$$(3.13.3) \quad \overline{W} \text{ is a linear subspace of } \ell^\infty(X, \mathbf{R}) \text{ or } \ell^\infty(X, \mathbf{C}),$$

as appropriate.

Similarly, if W is a subalgebra of $\ell^\infty(X, \mathbf{R})$ or $\ell^\infty(X, \mathbf{C})$, then

$$(3.13.4) \quad \overline{W} \text{ is a subalgebra of } \ell^\infty(X, \mathbf{R}) \text{ or } \ell^\infty(X, \mathbf{C}),$$

as appropriate. This corresponds to Theorem 7.29 on p161 of [210]. This uses the fact that if $\{f_j\}_{j=1}^\infty$ and $\{g_j\}_{j=1}^\infty$ are sequences of bounded real or complex-valued functions on X that converge uniformly to real or complex-valued functions f and g on X , respectively, then $\{f_j g_j\}_{j=1}^\infty$ converges uniformly to $f g$ on X . In order to show this, it is helpful to observe that $\{f_j\}_{j=1}^\infty$ and $\{g_j\}_{j=1}^\infty$ are uniformly bounded on X under these conditions.

3.13.2 Some more remarks about closures

Suppose now that $(X, d(x, y))$ is a nonempty metric space. The spaces $C_b(X, \mathbf{R})$ and $C_b(X, \mathbf{C})$ of bounded continuous real and complex-valued functions on X are subalgebras of $\ell^\infty(X, \mathbf{R})$ and $\ell^\infty(X, \mathbf{C})$, respectively. Remember that $C_b(X, \mathbf{R})$ and $C_b(X, \mathbf{C})$ are also closed sets in $\ell^\infty(X, \mathbf{R})$ and $\ell^\infty(X, \mathbf{C})$, respectively, with respect to the supremum metric, as in Subsection 1.14.1.

Let \mathcal{A} be a subalgebra of $C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$, so that \mathcal{A} may also be considered as a subalgebra of $\ell^\infty(X, \mathbf{R})$ or $\ell^\infty(X, \mathbf{C})$, as appropriate. The closure $\overline{\mathcal{A}}$ of \mathcal{A} in $C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$, as appropriate, is the same as the closure of \mathcal{A} in $\ell^\infty(X, \mathbf{R})$ or $\ell^\infty(X, \mathbf{C})$, as appropriate. In particular, $\overline{\mathcal{A}}$ is a subalgebra of $C_b(X, \mathbf{R})$ or $C_b(X, \mathbf{C})$, as appropriate, as before.

If X is compact, then one can use these remarks to reduce to the case of closed subalgebras \mathcal{A} of $C(X, \mathbf{R})$ or $C(X, \mathbf{C})$, with respect to the supremum metric, in the Stone–Weierstrass theorem. Of course, in this case, the statement that \mathcal{A} is dense in $C(X, \mathbf{R})$ or $C(X, \mathbf{C})$, as appropriate, with respect to the supremum metric, is the same as saying that \mathcal{A} is $C(X, \mathbf{R})$ or $C(X, \mathbf{C})$, as appropriate.

3.14 Some remarks about closed subalgebras

Let a be a positive real number. It is well known that the absolute value function on \mathbf{R} can be uniformly approximated on $[-a, a]$ by polynomials p with real coefficients such that

$$(3.14.1) \quad p(0) = 0,$$

as in Corollary 7.27 on p161 of [210]. Of course, Weierstrass' approximation theorem implies that $|\cdot|$ may be approximated uniformly on $[-a, a]$ by polynomials q with real coefficients. The constant term of such an approximation should be small, because $|0| = 0$, and so one can replace q with $q - q(0)$ to get approximations with the additional condition (3.14.1). This type of approximation can also be obtained more directly, without using Weierstrass' theorem, as in Exercise 23 on p169 of [210].

3.14.1 Closed subalgebras of $\ell^\infty(X, \mathbf{R})$

Let X be a nonempty set, and let \mathcal{A} be a subalgebra of the algebra $\ell^\infty(X, \mathbf{R})$ of all bounded real-valued functions on X . Suppose that \mathcal{A} is a closed set in $\ell^\infty(X, \mathbf{R})$, with respect to the supremum metric. If $f \in \mathcal{A}$, then it is well known that

$$(3.14.2) \quad |f| \in \mathcal{A}$$

too. This corresponds to Step 1 in the proof of Theorem 7.32 on p162 of [210], which is the Stone–Weierstrass theorem.

To see this, note that there is a nonnegative real number a such that $|f| \leq a$ on X , which means that

$$(3.14.3) \quad f(X) \subseteq [-a, a].$$

The absolute value function on \mathbf{R} can be uniformly approximated on $[-a, a]$ by polynomials p with real coefficients that satisfy (3.14.1), as before. It follows that $|f|$ can be uniformly approximated on X by the corresponding functions $p \circ f$. Remember that $p \circ f \in \mathcal{A}$ when (3.14.1) holds, as in Subsection 3.11.3. This implies (3.14.2), because \mathcal{A} is a closed set in $\ell^\infty(X, \mathbf{R})$, by hypothesis.

If g is another element of \mathcal{A} , then it is well known that

$$(3.14.4) \quad \max(f, g), \min(f, g) \in \mathcal{A}$$

as well, as in Step 2 on p163 of [210]. This is because

$$(3.14.5) \quad \max(f, g) = \frac{f + g}{2} + \frac{|f - g|}{2}$$

and

$$(3.14.6) \quad \min(f, g) = \frac{f + g}{2} - \frac{|f - g|}{2},$$

on X , as in [210]. Equivalently,

$$(3.14.7) \quad \max(f, g) + \min(f, g) = f + g$$

and

$$(3.14.8) \quad \max(f, g) - \min(f, g) = |f - g|$$

on X .

3.14.2 Closed subalgebras of $C(X, \mathbf{R})$

Suppose now that $(X, d(\cdot, \cdot))$ is a nonempty compact metric space, and that \mathcal{A} is a closed subalgebra of $C(X, \mathbf{R})$ with respect to the supremum metric. Suppose also that \mathcal{A} is nowhere vanishing on X , and that \mathcal{A} separates points in X . We would like to show that

$$(3.14.9) \quad \mathcal{A} = C(X, \mathbf{R}),$$

as in the Stone–Weierstrass theorem.

Let f be a continuous real-valued function on X , and let x, y be elements of X . As in Subsection 3.11.2, there is an element $h_{x,y}$ of \mathcal{A} such that

$$(3.14.10) \quad h_{x,y}(x) = f(x), \quad h_{x,y}(y) = f(y).$$

Let $\epsilon > 0$ be given. Because f and $h_{x,y}$ are continuous on X , there is an open subset $U_{x,y}$ of X such that $y \in U_{x,y}$ and

$$(3.14.11) \quad h_{x,y} > f - \epsilon \text{ on } U_{x,y}.$$

The collection of open sets $U_{x,y}$ of this type, $y \in X$, is an open covering of X . Using the compactness of X , we get that there are finitely many elements y_1, \dots, y_n of X such that

$$(3.14.12) \quad X \subseteq \bigcup_{j=1}^n U_{x,y_j}.$$

Put

$$(3.14.13) \quad g_x = \max(h_{x,y_1}, \dots, h_{x,y_n}).$$

Note that

$$(3.14.14) \quad g_x(x) = f(x),$$

by (3.14.10), and that

$$(3.14.15) \quad g_x > f - \epsilon \text{ on } X,$$

by (3.14.11) and (3.14.12). We also have that

$$(3.14.16) \quad g_x \in \mathcal{A},$$

as in the previous subsection. This corresponds to Step 3 on p163 of [210].

3.14.3 Using the g_x 's

Because f and g_x are continuous on X , there is an open subset V_x of X such that $x \in V_x$ and

$$(3.14.17) \quad g_x < f + \epsilon \text{ on } V_x.$$

We can use the compactness of X again to get finitely many elements x_1, \dots, x_m of X such that

$$(3.14.18) \quad X \subseteq \bigcup_{l=1}^m V_{x_l}.$$

Put

$$(3.14.19) \quad h = \min(g_{x_1}, \dots, g_{x_m}).$$

Observe that

$$(3.14.20) \quad h > f - \epsilon$$

on X , by (3.14.15), and that

$$(3.14.21) \quad h < f + \epsilon$$

on X , by (3.14.17) and (3.14.18). We also have that

$$(3.14.22) \quad h \in \mathcal{A},$$

because of (3.14.16), as before.

It follows that

$$(3.14.23) \quad f \in \mathcal{A},$$

because \mathcal{A} is a closed subalgebra of $C(X, \mathbf{R})$, by hypothesis. This corresponds to Step 4 on p164 of [210].

Chapter 4

Compactness and completeness

4.1 Diameters of sets

Let $(X, d(x, y))$ be a metric space, and let E be a subset of X . If E is nonempty and bounded, then the *diameter* of E can be defined as a nonnegative real number by

$$(4.1.1) \quad \text{diam } E = \sup\{d(x, y) : x, y \in E\}.$$

More precisely, one can use the boundedness of E to get that the set of distances $d(x, y)$ with $x, y \in E$ has an upper bound in \mathbf{R} . Conversely, if this set has an upper bound in \mathbf{R} , then it is easy to see that E is bounded in X . It is sometimes convenient to define the diameter of the empty set to be 0, and to interpret the diameter of an unbounded set as being $+\infty$.

Remember that if E is a bounded subset of X , then the closure \overline{E} of E is bounded in X too, as mentioned in Subsection 1.9.3. One can check that

$$(4.1.2) \quad \text{diam } \overline{E} = \text{diam } E$$

in this case.

4.1.1 Some properties of the diameter

If $E_1 \subseteq E_2 \subseteq X$, then

$$(4.1.3) \quad \text{diam } E_1 \leq \text{diam } E_2.$$

Observe that

$$(4.1.4) \quad \text{diam } \overline{B}(x, r) \leq 2r$$

for every $x \in X$ and nonnegative real number r , by the triangle inequality. Here $\overline{B}(x, r)$ is the closed ball in X centered at x with radius r , as in Subsection 1.9.3.

If $E \subseteq X$ is a bounded set, then

$$(4.1.5) \quad E \subseteq \overline{B}(x, \text{diam } E)$$

for every $x \in E$.

4.1.2 Some nested intersections

Let E_1, E_2, E_3, \dots be an infinite sequence of nonempty subsets of X such that

$$(4.1.6) \quad E_j \text{ is closed and bounded}$$

for each $j \geq 1$,

$$(4.1.7) \quad E_{j+1} \subseteq E_j$$

for every $j \geq 1$, and

$$(4.1.8) \quad \lim_{j \rightarrow \infty} \text{diam } E_j = 0.$$

One can check that $\bigcap_{j=1}^{\infty} E_j$ can have at most one element, because of (4.1.8).

Let x_j be an element of E_j for each $j \geq 1$. If $1 \leq j \leq l$, then it follows that

$$(4.1.9) \quad x_l \in E_l \subseteq E_j.$$

This implies that

$$(4.1.10) \quad d(x_j, x_l) \leq \text{diam } E_j$$

when $j \leq l$. Using this, it is easy to see that

$$(4.1.11) \quad \{x_j\}_{j=1}^{\infty} \text{ is a Cauchy sequence in } X.$$

Suppose for the moment that X is complete with respect to d , so that

$$(4.1.12) \quad \{x_j\}_{j=1}^{\infty} \text{ converges to an element } x \text{ of } X.$$

If l is any positive integer, then $\{x_j\}_{j=l}^{\infty}$ is a sequence of elements of E_l that converges to x . This implies that

$$(4.1.13) \quad x \in E_l$$

for every $l \geq 1$, because E_l is a closed set, by hypothesis. Thus

$$(4.1.14) \quad \bigcap_{l=1}^{\infty} E_l \neq \emptyset$$

under these conditions.

4.1.3 Necessity of completeness

Now let $\{x_j\}_{j=1}^\infty$ be any sequence of elements of X , and put

$$(4.1.15) \quad A_j = \{x_l : l \geq j\}$$

for each $j \geq 1$. Also put

$$(4.1.16) \quad E_j = \overline{A_j}$$

for each j . Of course, $A_j \neq \emptyset$ for each j , so that $E_j \neq \emptyset$. Note that

$$(4.1.17) \quad A_{j+1} \subseteq A_j$$

for each j , by construction, so that (4.1.7) holds.

If $\{x_j\}_{j=1}^\infty$ is a Cauchy sequence in X , then $\{x_j\}_{j=1}^\infty$ is bounded in X , as mentioned in Subsection 1.9.4. This implies that A_j is a bounded set in X for each j , so that E_j is bounded as well. In fact we have that

$$(4.1.18) \quad \text{diam } E_j = \text{diam } A_j \rightarrow 0 \text{ as } j \rightarrow \infty$$

under these conditions.

If

$$(4.1.19) \quad x \in \bigcap_{j=1}^\infty E_j,$$

then one can check that

$$(4.1.20) \quad \{x_j\}_{j=1}^\infty \text{ converges to } x \text{ in } X.$$

This uses the fact that

$$(4.1.21) \quad d(x, x_j) \leq \text{diam } E_j$$

for each j .

This shows that completeness of X is necessary in order to get (4.1.14) under the conditions described in the preceding subsection.

4.2 Totally bounded sets

Let $(X, d(x, y))$ be a metric space again. A subset E of X is said to be *totally bounded* if for every $r > 0$ there are finitely many elements x_1, \dots, x_l of X such that

$$(4.2.1) \quad E \subseteq \bigcup_{j=1}^l B(x_j, r).$$

If $E = \emptyset$, then this may be considered to hold trivially with $l = 0$ for each $r > 0$, even when $X = \emptyset$.

If (4.2.1) holds for some $r > 0$ and finitely many points x_1, \dots, x_l in X , then E is a bounded set in X , as in Section 1.9.2. Thus totally bounded sets are bounded in particular.

If $E \subseteq X$ is compact, then

$$(4.2.2) \quad E \text{ is totally bounded.}$$

More precisely, for each $r > 0$, E can be covered by open balls in X of radius r . Compactness implies that this open covering can be reduced to a finite subcovering, as desired.

Suppose that E is a bounded subset of \mathbf{R}^n for some positive integer n , with respect to the standard Euclidean metric on \mathbf{R}^n . In this case, one can check directly that E is totally bounded. This is simpler when $n = 1$, so that one can consider subsets of an interval in the real line. Similarly, a bounded set in \mathbf{R}^n is contained in a cube.

Of course, finite subsets of X are totally bounded. If X is equipped with the discrete metric and $E \subseteq X$ is totally bounded, then

$$(4.2.3) \quad E \text{ has only finitely many elements.}$$

This can be seen by taking $r = 1$ in (4.2.1).

4.2.1 Using closed balls

A subset E of X is totally bounded if and only if for every $t > 0$ there are finitely many elements y_1, \dots, y_l of X such that

$$(4.2.4) \quad E \subseteq \bigcup_{j=1}^l \overline{B}(y_j, t).$$

More precisely, the previous formulation implies this one, because $B(x, r)$ is contained in $\overline{B}(x, r)$ for every $x \in X$ and $r > 0$. To get the converse, one can use the fact that $\overline{B}(y, t)$ is contained in $B(y, r)$ when $t < r$.

If E is totally bounded, then it follows from this reformulation that

$$(4.2.5) \quad \text{the closure } \overline{E} \text{ of } E \text{ in } X \text{ is totally bounded}$$

too.

4.2.2 Using sets with small diameter

As another reformulation, a subset E of X is totally bounded if and only if for every $r_1 > 0$,

$$(4.2.6) \quad E \text{ is contained in the union of finitely many subsets of } X \\ \text{with diameter strictly less than } r_1.$$

More precisely, this formulation can be obtained from (4.2.1) using (4.1.4). Similarly, one can get (4.2.1) from this formulation using (4.1.5).

Equivalently, $E \subseteq X$ is totally bounded if and only if for every $t_1 > 0$,

$$(4.2.7) \quad E \text{ is contained in the union of finitely many subsets of } X \\ \text{with diameter less than or equal to } t_1.$$

4.2.3 Another characterization of total boundedness

Suppose that $E \subseteq X$ is not totally bounded in X . This means that there is an $r > 0$ such that E cannot be covered by finitely many open balls in X of radius r .

In particular, $E \neq \emptyset$, and we let x_1 be an element of E . By hypothesis, E is not contained in $B(x_1, r)$, and so there is an element x_2 of E that is not in $B(x_1, r)$.

If elements x_1, \dots, x_l of E have been chosen in this way for some positive integer l , then E is not contained in $\bigcup_{j=1}^l B(x_j, r)$, by hypothesis. This permits us to choose $x_{l+1} \in E$ such that x_{l+1} is not in $B(x_j, r)$ when $1 \leq j \leq l$.

Continuing in this way, we get an infinite sequence $\{x_j\}_{j=1}^{\infty}$ of elements of E such that x_l is not in $B(x_j, r)$ when $j < l$. Equivalently,

$$(4.2.8) \quad d(x_j, x_l) \geq r$$

when $j < l$.

Conversely, if there is a sequence of elements of E with these properties, then one can check that E is not totally bounded. More precisely, it is easy to see that an open ball in X of radius $r/2$ can contain at most one term in the sequence, because of (4.2.8) and the triangle inequality. This implies that E cannot be covered by finitely many open balls of radius $r/2$.

4.3 Separable metric spaces

A metric space $(X, d(x, y))$ is said to be *separable* if

$$(4.3.1) \quad \text{there is a dense subset of } X \text{ with only} \\ \text{finitely or countably many elements.}$$

It is well known that

$$(4.3.2) \quad \mathbf{R}^n \text{ is separable}$$

with respect to the standard Euclidean metric for each positive integer n , as in Exercise 22 at the end of Chapter 2 of [210]. More precisely, one can check that the set of points in \mathbf{R}^n with rational coordinates is a countable dense set in \mathbf{R}^n , as in the hint given in [210]. This is simpler when $n = 1$, as usual.

Note that X is automatically dense in itself, with respect to any metric $d(\cdot, \cdot)$. If X has only finitely or countably many elements, then it follows that X is separable with respect to $d(\cdot, \cdot)$.

If X is any set equipped with the discrete metric, then

$$(4.3.3) \quad X \text{ is the only dense subset of itself.}$$

Thus X is separable with respect to the discrete metric if and only if

$$(4.3.4) \quad X \text{ has only finitely or countably many elements.}$$

4.3.1 A characterization of separability

Let $(X, d(x, y))$ be any metric space again. One can check that X is separable if and only if for each $\epsilon > 0$ there is a subset $A(\epsilon)$ of X such that

$$(4.3.5) \quad A(\epsilon) \text{ has only finitely or countably many elements}$$

and

$$(4.3.6) \quad \bigcup_{x \in A(\epsilon)} B(x, \epsilon) = X.$$

Indeed, if X is separable, then there is a dense set $A \subseteq X$ with only finitely or countably many elements. In this case, one can take

$$(4.3.7) \quad A(\epsilon) = A$$

for every $\epsilon > 0$.

Conversely, if $A(1/j) \subseteq X$ satisfies (4.3.6) with $\epsilon = 1/j$ for each positive integer j , then it is easy to see that

$$(4.3.8) \quad A = \bigcup_{j=1}^{\infty} A(1/j)$$

is dense in X . If $A(1/j)$ also has only finitely or countably many elements for each $j \geq 1$, then (4.3.8) has only finitely or countably many elements as well.

Note that X is totally bounded if and only if for every $\epsilon > 0$ there is a finite set $A(\epsilon) \subseteq X$ such that (4.3.6) holds. This implies that

$$(4.3.9) \quad X \text{ is separable,}$$

as before.

4.3.2 The definition of a base for the topology of X

A collection \mathcal{B} of open subsets of X is said to be a *base for the topology of X* if for every $x \in X$ and $r > 0$ there is a $U \in \mathcal{B}$ such that $x \in U$ and

$$(4.3.10) \quad U \subseteq B(x, r).$$

Equivalently, this means that every nonempty open subset of X can be expressed as a union of elements of \mathcal{B} . One may consider the empty set to be the union of the empty collection of sets.

4.3.3 Separability and countable bases

Let E be a dense subset of X , and let $\mathcal{B}(E)$ be the collection of subsets of X of the form $B(y, 1/j)$, where $y \in E$ and $j \in \mathbf{Z}_+$. We would like to show that

$$(4.3.11) \quad \mathcal{B}(E) \text{ is a base for the topology of } X.$$

To do this, let $x \in X$ and $r > 0$ be given. Let j be a positive integer such that $2/j < r$. Because E is dense in X , there is a $y \in E$ such that $d(x, y) < 1/j$. This implies that $x \in B(y, 1/j)$, and one can also verify that

$$(4.3.12) \quad B(y, 1/j) \subseteq B(x, r),$$

using the triangle inequality, as desired.

If E has only finitely or countably many elements, then one can check that

$$(4.3.13) \quad \mathcal{B}(E) \text{ has only finitely or countably many elements.}$$

Conversely, if there is a base \mathcal{B} for the topology of X with only finitely or countably many elements, then X is separable. More precisely, one can choose an element from every nonempty element of \mathcal{B} to get a dense set in X with only finitely or countably many elements.

4.4 Lindelöf's theorem

Let $(X, d(x, y))$ be a metric space, and suppose that \mathcal{B} is a base for the topology of X with only finitely or countably many elements. Let $\{U_\alpha\}_{\alpha \in A}$ be a family of open subsets of X , indexed by a nonempty set A . Under these conditions, a theorem of Lindelöf says that there is a subset A_1 of A such that

$$(4.4.1) \quad A_1 \text{ has only finitely or countably many elements}$$

and

$$(4.4.2) \quad \bigcup_{\alpha \in A_1} U_\alpha = \bigcup_{\alpha \in A} U_\alpha.$$

4.4.1 The proof of Lindelöf's theorem

To see this, put

$$(4.4.3) \quad \mathcal{B}_\alpha = \{V \in \mathcal{B} : V \subseteq U_\alpha\}$$

for every $\alpha \in A$. Observe that

$$(4.4.4) \quad \bigcup_{V \in \mathcal{B}_\alpha} V = U_\alpha$$

for every $\alpha \in A$, because \mathcal{B} is a base for the topology of X , and U_α is an open set.

Put

$$(4.4.5) \quad \mathcal{B}' = \bigcup_{\alpha \in A} \mathcal{B}_\alpha.$$

Note that

$$(4.4.6) \quad \bigcup_{V \in \mathcal{B}'} V = \bigcup_{\alpha \in A} \left(\bigcup_{V \in \mathcal{B}_\alpha} V \right) = \bigcup_{\alpha \in A} U_\alpha.$$

If $V \in \mathcal{B}'$, then $V \in \mathcal{B}_\alpha$ for some $\alpha \in A$, and we let $\alpha(V)$ be an element of A with this property. Let

$$(4.4.7) \quad A_1 = \{\alpha(V) : V \in \mathcal{B}'\}$$

be the collection of elements of A that have been chosen in this way.

Thus

$$(4.4.8) \quad \bigcup_{V \in \mathcal{B}'} V \subseteq \bigcup_{\alpha \in A_1} U_\alpha,$$

because $V \subseteq U_{\alpha(V)}$ for every $V \in \mathcal{B}'$, by construction. This implies that

$$(4.4.9) \quad \bigcup_{\alpha \in A} U_\alpha \subseteq \bigcup_{\alpha \in A_1} U_\alpha,$$

by (4.4.6). The opposite inclusion holds automatically, because $A_1 \subseteq A$. It follows that (4.4.2) holds.

Clearly \mathcal{B}' has only finitely or countably many elements, because $\mathcal{B}' \subseteq \mathcal{B}$, and \mathcal{B} has only finitely or countably many elements, by hypothesis. It follows that A_1 has only finitely or countably many elements too, as desired.

4.4.2 The Lindelöf property

A subset E of X is said to have the *Lindelöf property* if every open covering of E in X can be reduced to a subcovering with only finitely or countably many elements. More precisely, this means that if $\{U_\alpha\}_{\alpha \in A}$ is an open covering of E in X , then there is a subset A_1 of A such that A_1 has only finitely or countably many elements and

$$(4.4.10) \quad E \subseteq \bigcup_{\alpha \in A_1} U_\alpha.$$

If there is a base \mathcal{B} for the topology of X with only finitely or countably many elements, then Lindelöf's theorem implies that every subset of X has the Lindelöf property.

Suppose that X has the Lindelöf property, and let $\epsilon > 0$ be given. Of course, X is covered by the collection of all open balls in X of radius ϵ . Because X has the Lindelöf property, there is a subset $A(\epsilon)$ of X with only finitely or countably many elements such that (4.3.6) holds. This implies that X is separable, as in Subsection 4.3.1.

4.5 The limit point property

Let $(X, d(x, y))$ be a metric space. A subset E of X is said to have the *limit point property* if for every infinite subset A of E there is an $x \in E$ such that

$$(4.5.1) \quad x \text{ is a limit point of } A \text{ in } X.$$

If E has only finitely many elements, then this condition holds vacuously.

4.5.1 Using compactness to get the limit point property

If E is compact, then it is well known that

$$(4.5.2) \quad E \text{ has the limit point property.}$$

To see this, let A be an infinite subset of E . Suppose for the sake of a contradiction that A does not have a limit point in E . This means that for every $x \in E$ there is a positive real number $r(x)$ such that $B(x, r(x))$ does not contain any element of A , except perhaps for x itself. Note that the collection of open balls of this type is an open covering of E in X .

Because E is compact, there are finitely many elements x_1, \dots, x_n of E such that

$$(4.5.3) \quad E \subseteq \bigcup_{j=1}^n B(x_j, r(x_j)).$$

In particular, this implies that

$$(4.5.4) \quad A \subseteq \bigcup_{j=1}^n (A \cap B(x_j, r(x_j))).$$

This means that A has at most n elements. This contradicts the hypothesis that A have infinitely many elements.

4.5.2 The limit point property implies total boundedness

If $E \subseteq X$ has the limit point property, then

$$(4.5.5) \quad E \text{ is totally bounded in } X.$$

Otherwise, there is an $r > 0$ and an infinite sequence $\{x_j\}_{j=1}^{\infty}$ of elements of X such that

$$(4.5.6) \quad d(x_j, x_l) \geq r \text{ when } j < l,$$

as in Subsection 4.2.3.

Let A be the set of x_j 's, $j \geq 1$. This is an infinite subset of E , because the x_j 's are distinct.

If $x \in X$, then it is easy to see that

$$(4.5.7) \quad A \cap B(x, r/2) \text{ has at most two elements,}$$

using the triangle inequality. This implies that x is not a limit point of A . It follows that E does not have the limit point property.

4.5.3 Some more remarks about the limit point property

Let Y be a subset of X , so that Y may be considered as a metric space too, using the restriction of $d(x, y)$ to $x, y \in Y$. If $E \subseteq Y$, then it is easy to see that

$$(4.5.8) \quad E \text{ has the limit point property as a subset of } X$$

if and only if

$$(4.5.9) \quad E \text{ has the limit point property as a subset of } Y.$$

In particular, one can take $Y = E$ here.

If $E \subseteq X$ has the limit point property, $E_0 \subseteq E$, and E_0 is a closed set in X , then one can check that

$$(4.5.10) \quad E_0 \text{ has the limit point property.}$$

If $E_1, E_2 \subseteq X$ have the limit point property, then one can verify that

$$(4.5.11) \quad E_1 \cup E_2 \text{ has the limit point property.}$$

4.6 Countable open coverings

Let $(X, d(x, y))$ be a metric space, and suppose that $E \subseteq X$ has the limit point property. Also let U_1, U_2, U_3, \dots be an infinite sequence of open subsets of X such that

$$(4.6.1) \quad E \subseteq \bigcup_{j=1}^{\infty} U_j.$$

We would like to show that there is a positive integer n such that

$$(4.6.2) \quad E \subseteq \bigcup_{j=1}^n U_j.$$

Otherwise, for each $n \geq 1$, we can choose $x_n \in E$ such that

$$(4.6.3) \quad x_n \notin \bigcup_{j=1}^n U_j.$$

Let A be the set of points $x_n, n \geq 1$, that have been chosen in this way. Let us check that

$$(4.6.4) \quad A \text{ has infinitely many elements.}$$

If $y \in E$, then $y \in U_l$ for some $l \geq 1$, by (4.6.1). This means that $x_n \neq y$ when $n \geq l$, because $x_n \notin U_l$, by construction. Thus $x_n = y$ for at most finitely many $n \geq 1$. This implies (4.6.4).

Hence there is an $x \in E$ such that x is a limit point of A in X , because E has the limit point property, by hypothesis. We also have that $x \in U_j$ for some $j \geq 1$, by (4.6.1). Because U_j is an open set in X , there is an $r > 0$ such that

$$(4.6.5) \quad B(x, r) \subseteq U_j.$$

There are infinitely many elements of A in $B(x, r)$, because x is a limit point of A . This means that there are infinitely many elements of A in U_j , by (4.6.5). Thus $x_n \in U_j$ for infinitely many $n \geq 1$. This contradicts the fact that $x_n \notin U_j$ when $n \geq j$, by construction.

4.6.1 Using the limit point property to get compactness

Suppose that X has the limit point property, and let us show that

$$(4.6.6) \quad X \text{ is compact.}$$

Remember that X is totally bounded in this case, as in Subsection 4.5.2. This implies that X is separable, as in Subsection 4.3.1.

It follows that there is a base for the topology of X with only finitely or countably many elements, as in Subsection 4.3.3. This means that X has the Lindelöf property, as in Subsection 4.4.2. The argument at the beginning of the section implies that any countable open covering of X can be reduced to a finite subcovering. Combining these two statements, we get that every open covering of X can be reduced to a finite subcovering.

4.6.2 Subsets of X with the limit point property

Now let E be any subset of X with the limit point property, and let us verify that

$$(4.6.7) \quad E \text{ is compact.}$$

Remember that E may also be considered as a metric space, with respect to the restriction of $d(x, y)$ to $x, y \in E$. We also have that E has the limit point property as a subset of itself, as in Subsection 4.5.3. This implies that E is compact as a subset of itself, as in the previous subsection. It is well known that this implies that E is compact as a subset of X , as desired.

4.7 Sequential compactness

Let $(X, d(x, y))$ be a metric space again. A subset E of X is said to be *sequentially compact* if for every sequence $\{x_j\}_{j=1}^{\infty}$ of elements of E there is a subsequence $\{x_{j_l}\}_{l=1}^{\infty}$ of $\{x_j\}_{j=1}^{\infty}$ and an element x of E such that

$$(4.7.1) \quad \{x_{j_l}\}_{l=1}^{\infty} \text{ converges to } x \text{ in } X.$$

4.7.1 The limit point property and sequential compactness

If E has the limit point property, then it is well known that

$$(4.7.2) \quad E \text{ is sequentially compact.}$$

To see this, let $\{x_j\}_{j=1}^{\infty}$ be a sequence of elements of E , and let A be the set of x_j 's, $j \geq 1$.

If A has only finitely many elements, then there is a subsequence $\{x_{j_l}\}_{l=1}^{\infty}$ of $\{x_j\}_{j=1}^{\infty}$ and an element x of E such that $x_{j_l} = x$ for every $l \geq 1$. In particular, this implies that $\{x_{j_l}\}_{l=1}^{\infty}$ converges to x in X .

Otherwise, if A has infinitely many elements, then there is an $x \in E$ such that x is a limit point of A in X , because E has the limit point property. This implies that for every positive real number r , there are infinitely many elements of A in $B(x, r)$, by a well-known property of limit points. This means that

$$(4.7.3) \quad d(x, x_j) < r$$

for infinitely many j , because of the way that A is defined.

In this case, one can show that there is a subsequence $\{x_{j_l}\}_{l=1}^{\infty}$ of $\{x_j\}_{j=1}^{\infty}$ that converges to x in X . In fact, one can choose $\{x_{j_l}\}_{l=1}^{\infty}$ so that

$$(4.7.4) \quad d(x, x_{j_l}) < 1/l$$

for each l .

4.7.2 The converse

Conversely, if $E \subseteq X$ is sequentially compact, then E has the limit point property. Indeed, let A be an infinite subset of E , and let $\{x_j\}_{j=1}^{\infty}$ be a sequence of distinct elements of A . By hypothesis, there is a subsequence $\{x_{j_l}\}_{l=1}^{\infty}$ of $\{x_j\}_{j=1}^{\infty}$ that converges to an element x of E .

One can check that x is a limit point of A in X under these conditions. More precisely, for each $r > 0$, $B(x, r)$ contains x_{j_l} for all but finitely many l , so that $B(x, r)$ contains infinitely many elements of A in particular.

4.7.3 More on sequential compactness

Let Y be a subset of X , so that Y may be considered as a metric space with respect to the restriction of the metric on X to elements of Y , and let E be a subset of Y . One can check directly that

$$(4.7.5) \quad E \text{ is sequentially compact as a subset of } X$$

if and only if

$$(4.7.6) \quad E \text{ is sequentially compact as a subset of } Y.$$

In particular, one can take $Y = E$ here, as usual.

If $E \subseteq X$ is sequentially compact, $E_0 \subseteq E$, and E_0 is a closed set in X , then one can verify directly that

$$(4.7.7) \quad E_0 \text{ is sequentially compact.}$$

If $E_1, E_2 \subseteq X$ are sequentially compact, then one can also verify directly that

$$(4.7.8) \quad E_1 \cup E_2 \text{ is sequentially compact.}$$

These statements could be obtained from their analogues in Subsection 4.5.3 for the limit point property and the equivalence of sequential compactness with the limit point property. However, it is interesting to consider the direct arguments as well. Of course, there are well-known arguments for the analogous statements for compactness in terms of open coverings too.

4.8 The Cauchy subsequence property

Let $(X, d(x, y))$ be a metric space. Let us say that $E \subseteq X$ has the *Cauchy subsequence property* if every sequence $\{x_j\}_{j=1}^\infty$ of elements of E has a subsequence $\{x_{j_l}\}_{l=1}^\infty$ that is a Cauchy sequence in X . Sequential compactness automatically implies the Cauchy subsequence property, because convergent sequences are Cauchy sequences.

If X is complete, E is a closed set in X , and E has the Cauchy subsequence property, then it is easy to see that E is sequentially compact.

We shall see that $E \subseteq X$ has the Cauchy subsequence property if and only if E is totally bounded. This corresponds to Theorem 6.3 H on p140 of [93], and to the equivalence of (b) and (c) in Theorem 74 on p101 of [130]. Theorem 5.6 on p22 of [81] states that X has the Cauchy subsequence property when X is totally bounded. It is not difficult to reduce to the case where $E = X$, and the fact that the Cauchy subsequence property implies total boundedness is somewhat easier to show.

4.8.1 The small subsequence property

Let us say that a sequence $\{x_j\}_{j=1}^\infty$ of elements of X is an ϵ -sequence for some $\epsilon > 0$ if

$$(4.8.1) \quad d(x_j, x_l) < \epsilon$$

for every $j, l \geq 1$. Let us say that a subset E of X has the *small subsequence property* if for every sequence $\{x_j\}_{j=1}^\infty$ of elements of E and every $\epsilon > 0$ there is a subsequence $\{x_{j_l}\}_{l=1}^\infty$ of $\{x_j\}_{j=1}^\infty$ such that

$$(4.8.2) \quad \{x_{j_l}\}_{l=1}^\infty \text{ is an } \epsilon\text{-sequence.}$$

If E has the Cauchy subsequence property, then it is easy to see that E has the small subsequence property. This uses the fact that every Cauchy sequence in X is an ϵ -sequence for any $\epsilon > 0$ after skipping finitely many terms in the sequence.

4.8.2 Total boundedness and small subsequences

If $E \subseteq X$ is totally bounded, then E has the small subsequence property. To see this, let a sequence $\{x_j\}_{j=1}^\infty$ of elements of E and an $\epsilon > 0$ be given. Because E is totally bounded, E can be covered by finitely many open balls in X of radius $\epsilon/2$. This implies that there is a single open ball of radius $\epsilon/2$ that contains x_j for infinitely many $j \geq 1$.

It follows that there is a subsequence $\{x_{j_l}\}_{l=1}^\infty$ of $\{x_j\}_{j=1}^\infty$ such that all of the x_{j_l} 's are contained in that open ball of radius $\epsilon/2$. This means that $\{x_{j_l}\}_{l=1}^\infty$ is an ϵ -sequence, as desired, by the triangle inequality.

Conversely, suppose that $E \subseteq X$ has the small subsequence property, and let us check that E is totally bounded. Otherwise, if E is not totally bounded, then there is an $r > 0$ and a sequence $\{x_j\}_{j=1}^\infty$ of elements of E such that

$d(x_j, x_l) \geq r$ when $j < l$, as in Subsection 4.2.3. In this case, $\{x_j\}_{j=1}^\infty$ has no subsequence that is an r -subsequence, so that E does not have the small subsequence property. In particular, if E has the Cauchy subsequence property, then E is totally bounded.

4.8.3 Small subsequences and Cauchy subsequences

If $E \subseteq X$ has the small subsequence property, then E has the Cauchy subsequence property. To see this, let a sequence $\{x_j\}_{j=1}^\infty$ of elements of E be given. By hypothesis, there is a subsequence of $\{x_j\}_{j=1}^\infty$ that is a $(1/2)$ -sequence.

We can repeat the process, to get a subsequence of the first subsequence that is a $(1/4)$ -sequence. Continuing in this way, we get an infinite sequence of subsequences of $\{x_j\}_{j=1}^\infty$, such that the $(n+1)$ th subsequence of $\{x_j\}_{j=1}^\infty$ is also a subsequence of the n th subsequence of $\{x_j\}_{j=1}^\infty$ for every $n \geq 1$, and the n th subsequence is a 2^{-n} -sequence for every $n \geq 1$.

One can check that the sequence obtained by taking the r th term of the r th subsequence for each $r \geq 1$ is a subsequence of $\{x_j\}_{j=1}^\infty$ as well. This subsequence is a Cauchy sequence too, because the terms with $r \geq n$ may be considered as a subsequence of the n th subsequence for each $n \geq 1$.

If E is totally bounded, then it follows that E has the Cauchy subsequence property.

4.9 A criterion for compactness

Let $(X, d(x, y))$ be a complete metric space. If

$$(4.9.1) \quad E \subseteq X \text{ is closed and totally bounded,}$$

then it is well known that E is compact. This corresponds to the fact that (5.3) implies (5.1) in Theorem 5.1 on p20 of [81] when $E = X$, and it is not difficult to reduce to that case. This also corresponds to the fact that (c) implies (a) in Theorem A 4 on p369 of [208].

This basically corresponds to the fact that (c) implies (a) in Theorem 74 on p101 of [130] as well. Note that compactness of a subset of a metric space is defined on p99 of [130] in terms of the limit point property or equivalently sequential compactness. The equivalence of this with the definition of compactness in terms of open coverings is part of Theorem 72 on p101 of [130].

Compactness of a metric space is defined on p145 of [93] to mean that the metric space is complete and totally bounded. The definition of compactness in terms of open coverings is called the Heine–Borel property in [93], and the equivalence of these conditions is established just afterwards.

4.9.1 Using the Cauchy subsequence property

One way to show this is to use the fact that E has the Cauchy subsequence property, because E is totally bounded, as in the previous section. In this case,

Cauchy sequences converge in X , because X is complete. We also have that the limit of a convergent sequence of elements of E is contained in E , because E is a closed set. It follows that E is sequentially compact under these conditions.

This implies that E has the limit point property, as in Subsection 4.7.1, so that E is compact, as in Subsection 4.6.1.

4.9.2 A reformulation of total boundedness

Before describing another proof, let us reformulate the hypothesis that E be totally bounded. If r is a positive real number, then E is contained in the union of finitely many subsets of X with diameter less than or equal to r , as in Subsection 4.2.2. We may as well suppose that these are subsets of E , since we can replace them with their intersections with E .

We can also take these subsets to be closed sets in X , by replacing them with their closures, which have the same diameters, as in Section 4.1. The closures of these subsets of E are contained in E , because E is a closed set, by hypothesis.

4.9.3 Another proof of compactness

Let $\{U_\alpha\}_{\alpha \in A}$ be an open covering of E in X for which there is no finite sub-covering, so that $E \neq \emptyset$ in particular. As in the previous subsection, E can be expressed as the union of finitely many closed sets, each of which has diameter less than or equal to $1/2$. At least one of these finitely many sets cannot be covered by finitely many U_α 's. This implies that there is a subset E_1 of E such that E_1 is a closed set in X , the diameter of E_1 is less than or equal to $1/2$, and E_1 cannot be covered by finitely many U_α 's.

Continuing in this way, we can get an infinite sequence E_1, E_2, E_3, \dots of subsets of E such that E_j is a closed set in X for every $j \geq 1$,

$$(4.9.2) \quad E_{j+1} \subseteq E_j$$

for every $j \geq 1$,

$$(4.9.3) \quad \text{diam } E_j \leq 2^{-j}$$

for every $j \geq 1$, and E_j cannot be covered by finitely many U_α 's for any $j \geq 1$. More precisely, suppose that E_j has been chosen for some $j \geq 1$, and let us choose E_{j+1} . By hypothesis, E_j is a closed set contained in E , and hence E_j is totally bounded. Thus E_j can be expressed as the union of finitely many closed sets, each of which has diameter less than or equal to 2^{-j-1} . At least one of these subsets cannot be covered by finitely many U_α 's, because E_j cannot be covered by finitely many U_α 's. Thus we take E_{j+1} to be one of these subsets of E_j , in such a way that E_{j+1} cannot be covered by finitely many U_α 's. Clearly E_{j+1} satisfies (4.9.2), and the analogues of the other conditions for $j+1$.

Note that $\text{diam } E_j \rightarrow 0$ as $j \rightarrow \infty$, by (4.9.3). This implies that there is an $x \in X$ that is contained in E_j for every $j \geq 1$, because X is complete, as in Subsection 4.1.2.

In particular, $x \in E$, so that there is an $\alpha_0 \in A$ such that $x \in U_{\alpha_0}$. It follows that there is a positive real number r such that

$$(4.9.4) \quad B(x, r) \subseteq U_{\alpha_0},$$

because U_{α_0} is an open set. If j is large enough so that $2^{-j} < r$, then E_j is contained in $B(x, r)$, because of (4.9.3) and the fact that $x \in E_j$. This means that $E_j \subseteq U_{\alpha_0}$, by (4.9.4). This contradicts the condition that E_j cannot be covered by finitely many U_α 's, as desired.

4.10 Cauchy sequences and convergent subsequences

Let $\{x_j\}_{j=1}^\infty$ be a Cauchy sequence of elements of a metric space (X, d) , which is not necessarily complete. If there is a subsequence $\{x_{j_l}\}_{l=1}^\infty$ of $\{x_j\}_{j=1}^\infty$ that converges to an element x of X , then it is well known and not too difficult to check that

$$(4.10.1) \quad \{x_j\}_{j=1}^\infty \text{ converges to } x$$

in X . Indeed, if ϵ is any positive real number, then there are positive integers L and N such that

$$(4.10.2) \quad d(x, x_{j_l}) < \epsilon/2$$

for all $l \geq L$, and

$$(4.10.3) \quad d(x_m, x_n) < \epsilon/2$$

for all $m, n \geq N$, by hypothesis. If $l \geq \max(L, N)$, then $j_l \geq l \geq N$, and we can use the triangle inequality to get that

$$(4.10.4) \quad d(x, x_n) \leq d(x, x_{j_l}) + d(x_{j_l}, x_n) < \epsilon/2 + \epsilon/2 = \epsilon$$

when $n \geq N$.

In particular, if $\{x_j\}_{j=1}^\infty$ is a sequence of elements of a sequentially compact set $E \subseteq X$, then $\{x_j\}_{j=1}^\infty$ converges to an element of E . If X is sequentially compact, then it follows that X is complete. Remember that compact subsets of metric spaces have the limit point property, and hence are sequentially compact.

4.10.1 Some additional remarks about compactness

We can also use the second proof of the criterion for compactness in the previous section as another way to show that sequential compactness implies compactness. More precisely, if a metric space X is sequentially compact, then X is complete and totally bounded, and hence compact, by this criterion.

Now let E be a subset of an arbitrary metric space X , and remember that E may be considered as a metric space as well, by restricting the metric to elements of E . If E is sequentially compact in X , then E is sequentially compact as a subset of itself, as in Subsection 4.7.3. This implies that E is compact as a subset of itself, as before, and hence that E is compact as a subset of X .

4.11 The Baire category theorem

Let $(X, d(x, y))$ be a metric space. One can check that $E \subseteq X$ is dense in X if and only if for every open set $V \subseteq X$ with $V \neq \emptyset$, we have that

$$(4.11.1) \quad E \cap V \neq \emptyset.$$

More precisely, if $x \in V$, then x is an element of E or a limit point of E , because E is dense in X , and (4.11.1) holds in both cases.

If E is dense in X and U is a dense open set in X , then one can verify that

$$(4.11.2) \quad E \cap U \text{ is dense in } X.$$

Indeed, if V is a nonempty open set in X , then it suffices to show that

$$(4.11.3) \quad E \cap U \cap V \neq \emptyset.$$

Note that $U \cap V \neq \emptyset$, because U is dense in X , by hypothesis, as in the preceding paragraph. Of course, $U \cap V$ is an open set in X too, so that (4.11.3) follows from the hypothesis that E be dense in X , as in the preceding paragraph again.

This implies that the intersection of two dense open sets in X is a dense open set in X as well. It follows that the intersection of finitely many dense open subsets of X is a dense open subset of X too.

4.11.1 Some infinite intersections

Let U_1, U_2, U_3, \dots be an infinite sequence of dense open subsets of X . If X is complete, then the *Baire category theorem* states that

$$(4.11.4) \quad \bigcap_{j=1}^{\infty} U_j \text{ is dense in } X.$$

This corresponds to Exercise 22 on p82 of [210], and to Exercise 30 on p46 of [210] for subsets of Euclidean spaces. This also corresponds to Theorem 2.6 on p11 of [81], and to Exercise 6 on p113 of [130]. If X is the real line with the standard Euclidean metric, then this corresponds to Exercise 5 on p131 of [93].

4.11.2 The beginning of the proof

To see this, let $x \in X$ and $r > 0$ be given, and let us show that

$$(4.11.5) \quad \overline{B}(x, r) \cap \left(\bigcap_{j=1}^{\infty} U_j \right) \neq \emptyset.$$

Because U_1 is dense in X , there is a $y_1 \in U_1$ such that $d(x, y_1) < r$. Let us choose $r_1 > 0$ small enough so that $r_1 \leq 1$,

$$(4.11.6) \quad \overline{B}(y_1, r_1) \subseteq U_1,$$

and

$$(4.11.7) \quad d(x, y_1) + r_1 \leq r.$$

This uses the hypothesis that U_1 be an open set to get (4.11.6). Note that

$$(4.11.8) \quad B(y_1, r_1) \subseteq B(x, r) \text{ and } \overline{B}(y_1, r_1) \subseteq \overline{B}(x, r),$$

because of (4.11.7) and the triangle inequality.

4.11.3 Repeating the process

We can repeat the process to get an infinite sequence $\{y_j\}_{j=1}^{\infty}$ of elements of X and an infinite sequence $\{r_j\}_{j=1}^{\infty}$ of positive real numbers with the following properties. First,

$$(4.11.9) \quad r_j \leq 1/j$$

for each $j \geq 1$. Second,

$$(4.11.10) \quad \overline{B}(y_j, r_j) \subseteq U_j$$

for every $j \geq 1$. Third, (4.11.7) holds when $j = 1$, and otherwise

$$(4.11.11) \quad d(y_{j-1}, y_j) + r_j \leq r_{j-1}$$

when $j \geq 2$. This implies that

$$(4.11.12) \quad B(y_j, r_j) \subseteq B(y_{j-1}, r_{j-1}) \text{ and } \overline{B}(y_j, r_j) \subseteq \overline{B}(y_{j-1}, r_{j-1})$$

when $j \geq 2$, using the triangle inequality again.

More precisely, we can do this when $j = 1$, as before. Suppose that $y_j \in X$ and $r_j > 0$ have been chosen with these properties for some $j \geq 1$, and let us see how we can choose y_{j+1} and r_{j+1} . Because U_{j+1} is dense in X , there is a $y_{j+1} \in U_{j+1}$ such that

$$(4.11.13) \quad d(y_j, y_{j+1}) < r_j.$$

We can choose $r_{j+1} > 0$ small enough so that $r_{j+1} \leq 1/(j+1)$,

$$(4.11.14) \quad \overline{B}(y_{j+1}, r_{j+1}) \subseteq U_{j+1},$$

and

$$(4.11.15) \quad d(y_j, y_{j+1}) + r_{j+1} \leq r_j.$$

This uses the hypothesis that U_{j+1} be an open set to get (4.11.14), as before.

4.11.4 The rest of the proof

If $1 \leq j \leq l$, then

$$(4.11.16) \quad B(y_l, r_l) \subseteq B(y_j, r_j) \text{ and } \overline{B}(y_l, r_l) \subseteq \overline{B}(y_j, r_j),$$

because of (4.11.12). In particular,

$$(4.11.17) \quad B(y_l, r_l) \subseteq B(x, r) \text{ and } \overline{B}(y_l, r_l) \subseteq \overline{B}(x, r)$$

for every $l \geq 1$, because of (4.11.8). Note that

$$(4.11.18) \quad d(y_j, y_l) < r_j$$

when $j \leq l$, by (4.11.16). This implies that

$$(4.11.19) \quad \{y_l\}_{l=1}^{\infty} \text{ is a Cauchy sequence in } X,$$

because of (4.11.9). It follows that $\{y_l\}_{l=1}^{\infty}$ converges to an element y of X , because X is complete.

Observe that $y_l \in \overline{B}(x, r)$ for every $l \geq 1$, by (4.11.17), so that

$$(4.11.20) \quad y \in \overline{B}(x, r).$$

Similarly, for each $j \geq 1$, we have that $y_l \in \overline{B}(y_j, r_j)$ when $l \geq j$, by (4.11.16). This implies that

$$(4.11.21) \quad y \in \overline{B}(y_j, r_j)$$

for every $j \geq 1$. It follows that

$$(4.11.22) \quad y \in U_j$$

for every $j \geq 1$, because of (4.11.10). Thus

$$(4.11.23) \quad y \in \overline{B}(x, r) \cap \left(\bigcap_{j=1}^{\infty} U_j \right),$$

so that (4.11.5) holds.

Because (4.11.5) holds for every $r > 0$, we get that x is an element of $\bigcap_{j=1}^{\infty} U_j$, or a limit point of $\bigcap_{j=1}^{\infty} U_j$. This implies (4.11.4), because $x \in X$ is arbitrary.

Some related results will be mentioned in the next section.

4.12 The interior of a set

Let $(X, d(\cdot, \cdot))$ be a metric space. The *interior* E° of a subset E of X is the set of $x \in E$ for which there is an $r > 0$ such that $B(x, r) \subseteq E$.

It is easy to see that

$$(4.12.1) \quad X \setminus E^\circ = \overline{X \setminus E}.$$

Note that (4.12.1) implies that E° is an open set in X , because the left side is a closed set in X . One can also verify that E° is an open set directly from the definitions.

Equivalently, if A is any subset of X , then

$$(4.12.2) \quad X \setminus \overline{A} = (X \setminus A)^\circ.$$

This can be used to show that \overline{A} is a closed set in X , by checking directly that the interior of any subset of X is an open set.

In particular, (4.12.1) implies that

$$(4.12.3) \quad E^\circ = \emptyset \text{ if and only if } X \setminus E \text{ is dense in } X.$$

4.12.1 Some finite unions of closed sets in X

Let E_1, \dots, E_n be finitely many closed sets in X , so that their union $\bigcup_{j=1}^n E_j$ is a closed set in X too. If $E_j^\circ = \emptyset$ for each $j = 1, \dots, n$, then we would like to check that

$$(4.12.4) \quad \left(\bigcup_{j=1}^n E_j \right)^\circ = \emptyset.$$

This is the same as saying that

$$(4.12.5) \quad X \setminus \left(\bigcup_{j=1}^n E_j \right) = \bigcap_{j=1}^n (X \setminus E_j)$$

is dense in X , as in (4.12.3).

Note that $X \setminus E_j$ is a dense open set in X for each j , by hypothesis. Thus we can use the fact that the intersection of finitely many dense open subsets of X is a dense open set in X as well, as in the previous section.

4.12.2 A reformulation of the Baire category theorem

Let E_1, E_2, E_3, \dots be an infinite sequence of closed subsets of X such that $E_j^\circ = \emptyset$ for every $j \geq 1$. The Baire category theorem is equivalent to saying that if X is complete, then

$$(4.12.6) \quad \left(\bigcup_{j=1}^{\infty} E_j \right)^\circ = \emptyset.$$

Indeed, this is the same as saying that

$$(4.12.7) \quad X \setminus \left(\bigcup_{j=1}^{\infty} E_j \right) = \bigcap_{j=1}^{\infty} (X \setminus E_j)$$

is dense in X , as in (4.12.3).

Note that $X \setminus E_j$ is a dense open set in X for each j , by hypothesis, and in fact these may be any dense open sets in X . The Baire category theorem is the same as saying that (4.12.7) is dense in X when X is complete, as in (4.11.4).

4.12.3 First and second category

A subset A of X is said to be *nowhere dense* in X if

$$(4.12.8) \quad \text{the closure } \overline{A} \text{ of } A \text{ in } X \text{ has empty interior.}$$

A subset B of X is said to be of *first category* in X , or *meager*, if

$$(4.12.9) \quad \begin{array}{l} B \text{ can be expressed as the union} \\ \text{of a sequence of nowhere dense sets.} \end{array}$$

Otherwise, B is said to be of *second category* in X , or *nonmeager*.

If $B \subseteq X$ is of first category and X is complete, then

$$(4.12.10) \quad B^\circ = \emptyset,$$

as in Corollary 2.7 on p11 of [81]. This can be obtained from the Baire category theorem, as in the previous subsection. More precisely, this uses the fact that B is contained in the union of a sequence of closed sets in X with empty interior, by hypothesis. In particular, if $X \neq \emptyset$, then

$$(4.12.11) \quad X \text{ is of second category,}$$

as in Theorem 79 on p112 of [130]. If X is the real line with the standard Euclidean metric, then this corresponds to Theorem 5.6 I on p130 of [93].

If B_1, B_2, B_3, \dots is an infinite sequence of subsets of X of first category, then it is not too difficult to show that

$$(4.12.12) \quad \bigcup_{l=1}^{\infty} B_l$$

is of first category in X as well.

Of course, if $x \in X$, then $\{x\}$ is a closed set in X . Suppose that every element of X is a limit point of X . This is the same as saying that if $x \in X$, then x is not an isolated point in X , so that $\{x\}$ has empty interior. If X is nonempty and complete, then one can use the Baire category theorem to get that X is uncountable.

Chapter 5

Equicontinuity and sequences of functions

Let $\{f_j\}_{j=1}^\infty$ be a sequence of mappings from a set X into a metric space Y . In some situations, we might like to find a subsequence $\{f_{j_l}\}_{l=1}^\infty$ of $\{f_j\}_{j=1}^\infty$ that converges to a mapping f from X into Y , at least in some sense.

If X is a metric space too, then we may be interested in additional continuity conditions on the f_j 's, and on f . In particular, X may be a subset of \mathbf{R}^n for some positive integer n . These and related matters will be discussed in this chapter.

5.1 Pointwise convergent subsequences

Let E be a nonempty set with only finitely or countably many elements, and let Y be a metric space. Also let $\{f_j\}_{j=1}^\infty$ be a sequence of mappings from E into Y . Suppose that for each $x \in E$ there is a sequentially compact set $K(x) \subseteq Y$ such that

$$(5.1.1) \quad f_j(x) \in K(x)$$

for every $j \geq 1$.

Under these conditions, we would like to show that there is a subsequence $\{f_{j_l}\}_{l=1}^\infty$ of $\{f_j\}_{j=1}^\infty$ that converges pointwise to a mapping f from E into Y , with

$$(5.1.2) \quad f(x) \in K(x)$$

for every $x \in E$. Remember that compact subsets of Y have the limit point property, as in Subsection 4.5.1, and thus are sequentially compact, as in Subsection 4.7.1.

5.1.1 Complex and \mathbf{R}^n -valued functions

Of course, (5.1.1) implies that

$$(5.1.3) \quad \{f_j(x)\}_{j=1}^\infty \text{ is a bounded sequence in } Y.$$

If Y is the complex plane, or \mathbf{R}^n for some positive integer n , with the standard Euclidean metric, and if $\{f_j(x)\}_{j=1}^\infty$ is a bounded sequence in Y , then there is a sequentially compact set $K(x) \subseteq Y$ such that (5.1.1) holds for each j . This uses the fact that subsets of \mathbf{C} or \mathbf{R}^n that are closed and bounded are compact, and thus sequentially compact.

5.1.2 Finite sets E

Suppose first that E has only finitely many elements x_1, \dots, x_n . Using sequential compactness of $K(x_1)$, we can get a subsequence $\{f_{j_i}\}_{i=1}^\infty$ of $\{f_j\}_{j=1}^\infty$ such that

$$(5.1.4) \quad \{f_{j_i}(x_1)\}_{i=1}^\infty \text{ converges to an element } f(x_1) \text{ of } K(x_1).$$

More precisely, there is a subsequence of $\{f_j(x_1)\}_{j=1}^\infty$ that converges to an element of $K(x_1)$, and we can use the same sequence of indices to consider this as a subsequence of $\{f_j\}_{j=1}^\infty$ that converges at x_1 to an element of $K(x_1)$.

If $n \geq 2$, then we can use sequential compactness of $K(x_2)$ to get a subsequence $\{f_{j_{i_r}}\}_{r=1}^\infty$ of $\{f_{j_i}\}_{i=1}^\infty$ such that

$$(5.1.5) \quad \{f_{j_{i_r}}(x_2)\}_{r=1}^\infty \text{ converges to an element } f(x_2) \text{ of } K(x_2).$$

As before, we use the sequential compactness of $K(x_2)$ to get a subsequence of $\{f_{j_i}(x_2)\}_{i=1}^\infty$ that converges to an element of $K(x_2)$, and we use the same sequence of indices to consider this as a subsequence of $\{f_{j_i}\}_{i=1}^\infty$ that converges at x_2 to an element of $K(x_2)$. Note that

$$(5.1.6) \quad \{f_{j_{i_r}}(x_1)\}_{r=1}^\infty \text{ converges to } f(x_1) \text{ in } Y,$$

because a subsequence of a convergent sequence converges to the same limit.

We can repeat the process until we get a subsequence of $\{f_j\}_{j=1}^\infty$ that converges pointwise on E in this case.

5.1.3 Countably infinite sets E

Suppose now that E is countably infinite, and let x_1, x_2, x_3, \dots be an enumeration of the elements of E . As before, we can use sequential compactness of $K(x_1)$ to get a subsequence of $\{f_j\}_{j=1}^\infty$ that converges pointwise at x_1 . Repeating the process, we get a sequence of subsequences of $\{f_j\}_{j=1}^\infty$ with the following properties.

First, for each positive integer r ,

$$(5.1.7) \quad \begin{array}{l} \text{the } r\text{th subsequence converges pointwise at } x_r \\ \text{to an element of } K(x_r). \end{array}$$

Second, if $r \geq 2$, then

(5.1.8) the r th subsequence is a subsequence of the $(r-1)$ th subsequence.

The second property ensures that the r th subsequence is a subsequence of all the previous subsequences, and of the initial sequence $\{f_j\}_{j=1}^\infty$. Combining this with the first property, we get that the r th subsequence converges pointwise at each of the previous points, with the same limits as the previous subsequences.

5.1.4 A single subsequence

Let f be the mapping from E into Y such that for each $r \geq 1$, $f(x_r)$ is the limit of the r th subsequence of $\{f_j\}_{j=1}^\infty$ at x_r . Thus

$$(5.1.9) \quad f(x_r) \in K(x_r)$$

for every $r \geq 1$, by construction. We would like to show that there is a subsequence of $\{f_j\}_{j=1}^\infty$ that converges to f pointwise on E .

Consider the sequence of mappings $\{g_n\}_{n=1}^\infty$ from E into Y obtained by taking g_n to be the n th term of the n th subsequence mentioned before for each $n \in \mathbf{Z}_+$. One can check that

$$(5.1.10) \quad \{g_n\}_{n=1}^\infty \text{ is a subsequence of } \{f_j\}_{j=1}^\infty.$$

More precisely, one can verify that for each positive integer $n+1$, g_{n+1} occurs after g_n in the original sequence of mappings. In fact, g_{n+1} occurs after the n th term in the n th subsequence.

Similarly, for each positive integer r ,

$$(5.1.11) \quad \{g_n\}_{n=r}^\infty \text{ is a subsequence of the } r\text{th subsequence mentioned earlier.}$$

This implies that

$$(5.1.12) \quad \{g_n(x_r)\}_{n=r}^\infty \text{ converges to } f(x_r)$$

in Y .

It follows that

$$(5.1.13) \quad \{g_n(x_r)\}_{n=1}^\infty \text{ converges to } f(x_r)$$

in Y for each $r \in \mathbf{Z}_+$, as desired.

5.2 Equicontinuity

Let (X, d_X) and (Y, d_Y) be metric spaces. A collection \mathcal{E} of mappings from X into Y is said to be *equicontinuous* at a point $x \in X$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$(5.2.1) \quad d_Y(f(x), f(w)) < \epsilon$$

for every $f \in \mathcal{E}$ and $w \in X$ with $d_X(x, w) < \delta$. In particular, this implies that every $f \in \mathcal{E}$ is continuous at x .

5.2.1 Uniform equicontinuity

Similarly, \mathcal{E} is said to be *uniformly equicontinuous* on X if for every $\epsilon > 0$ there is a $\delta > 0$ such that (5.2.1) holds for every $f \in \mathcal{E}$ and $x, w \in X$ with $d_X(x, w) < \delta$. This implies that \mathcal{E} is equicontinuous at every point in X , and that every element of \mathcal{E} is uniformly continuous on X .

If X is compact, then it is well known that every continuous mapping from X into Y is uniformly continuous. In this case, if \mathcal{E} is equicontinuous at every point in X , then one can show that

$$(5.2.2) \quad \mathcal{E} \text{ is uniformly equicontinuous on } X,$$

using essentially the same argument.

If $x \in X$ and every $f \in \mathcal{E}$ is continuous at x , then for each $\epsilon > 0$ and $f \in \mathcal{E}$ there is a $\delta_f(x, \epsilon) > 0$ such that (5.2.1) holds when $w \in X$ satisfies $d_X(x, w) < \delta_f(x, \epsilon)$. If \mathcal{E} has only finitely many elements, then

$$(5.2.3) \quad \min\{\delta_f(x, \epsilon) : f \in \mathcal{E}\} > 0,$$

and \mathcal{E} is equicontinuous at x . Similarly, if every element of \mathcal{E} is uniformly continuous on X , and \mathcal{E} has only finitely many elements, then \mathcal{E} is uniformly equicontinuous on X .

Let C be a nonnegative real number, and suppose that

$$(5.2.4) \quad \text{every element of } \mathcal{E} \text{ is Lipschitz on } X \text{ with constant } C,$$

as in Section 1.1. It is easy to see that \mathcal{E} is uniformly equicontinuous on X in this case. There is an analogous statement for Lipschitz or Hölder continuity conditions of any order $\alpha > 0$, as in Section A.2.

5.2.2 Equicontinuity and sequences of mappings

Let $\{f_j\}_{j=1}^{\infty}$ be a sequence of mappings from X into Y , and let x be an element of X . We say that $\{f_j\}_{j=1}^{\infty}$ is equicontinuous at x if the collection of f_j 's, $j \in \mathbf{Z}_+$, is equicontinuous at x .

Suppose that $\{f_j\}_{j=1}^{\infty}$ converges pointwise to a mapping f from X into Y , and that $\{f_j\}_{j=1}^{\infty}$ is equicontinuous at x . We would like to check that

$$(5.2.5) \quad f \text{ is continuous at } x$$

as well under these conditions.

Let $\epsilon > 0$ be given, so that there is a $\delta = \delta(x, \epsilon) > 0$ such that

$$(5.2.6) \quad d_Y(f_j(x), f_j(w)) < \epsilon$$

for every $j \in \mathbf{Z}_+$ and $w \in X$ with $d_X(x, w) < \delta$. If $w \in X$ satisfies $d_X(x, w) < \delta$, then one can check that

$$(5.2.7) \quad d_Y(f(x), f(w)) \leq \epsilon,$$

using (5.2.6) and the fact that $\{f_j(x)\}_{j=1}^\infty$ and $\{f_j(w)\}_{j=1}^\infty$ converge to $f(x)$ and $f(w)$ in Y , respectively, by hypothesis. More precisely,

$$\begin{aligned} d_Y(f(x), f(w)) &\leq d_Y(f(x), f_j(x)) + d_Y(f_j(x), f_j(w)) + d_Y(f_j(w), f(w)) \\ (5.2.8) \quad &< d_Y(f_j(x), f(x)) + \epsilon + d_Y(f_j(w), f(w)) \end{aligned}$$

for every $j \in \mathbf{Z}_+$ and $w \in X$ with $d_X(x, w) < \delta$, using the triangle inequality twice in the first step, and (5.2.6) in the second step. The right side is arbitrarily close to ϵ when j is sufficiently large, and one can use this to get (5.2.7). This implies that f is continuous at x , as desired.

5.2.3 Uniformly equicontinuous sequences of mappings

Similarly, a sequence $\{f_j\}_{j=1}^\infty$ of mappings from X into Y is said to be uniformly equicontinuous on X if the collection of f_j 's, $j \in \mathbf{Z}_+$, is uniformly equicontinuous on X . If $\{f_j\}_{j=1}^\infty$ is uniformly equicontinuous on X , and if $\{f_j\}_{j=1}^\infty$ converges pointwise to a mapping f from X into Y , then one can verify that

$$(5.2.9) \quad f \text{ is uniformly continuous on } X.$$

This is basically the same as the argument in the preceding subsection, except that $\delta = \delta(\epsilon)$ does not depend on $x \in X$.

5.2.4 Lipschitz conditions and pointwise convergence

Let C be a nonnegative real number again, and let $\{f_j\}_{j=1}^\infty$ be a sequence of mappings from X into Y such that f_j is Lipschitz on X with constant C for each j . If $\{f_j\}_{j=1}^\infty$ converges pointwise to a mapping f from X into Y , then one can check that

$$(5.2.10) \quad f \text{ is Lipschitz on } X \text{ with constant } C$$

too. There is an analogous statement for Lipschitz or Hölder continuity conditions of any order $\alpha > 0$, as before.

5.3 Uniformly Cauchy sequences

Let X be a set, and let (Y, d_Y) be a metric space. Let us say that a sequence $\{f_j\}_{j=1}^\infty$ of mappings from X into Y is *uniformly Cauchy* on X if for every $\epsilon > 0$ there is a positive integer $L(\epsilon)$ such that

$$(5.3.1) \quad d_Y(f_j(x), f_l(x)) < \epsilon$$

for every $x \in X$ and $j, l \geq L(\epsilon)$.

If $X \neq \emptyset$ and the f_j 's are bounded mappings from X into Y , then it is easy to see that this is equivalent to the condition that $\{f_j\}_{j=1}^\infty$ be a Cauchy sequence with respect to the supremum metric on the space of bounded mappings from X into Y . This is analogous to the equivalence between uniform convergence and convergence with respect to the supremum metric for bounded mappings from X into Y , as in Subsection 1.11.3.

5.3.1 Uniform convergence implies uniformly Cauchy

If $\{f_j\}_{j=1}^\infty$ is any sequence of mappings from X into Y that converges uniformly to a mapping f from X into Y , then one can check that

$$(5.3.2) \quad \{f_j\}_{j=1}^\infty \text{ is uniformly Cauchy on } X.$$

This is analogous to the fact that convergent sequences in a metric space are Cauchy sequences.

5.3.2 Convergence when Y is complete

Let $\{f_j\}_{j=1}^\infty$ be a uniformly Cauchy sequence of mappings from X into Y . In particular, this implies that

$$(5.3.3) \quad \{f_j(x)\}_{j=1}^\infty \text{ is a Cauchy sequence in } Y$$

for every $x \in X$. If Y is complete, then it follows that $\{f_j\}_{j=1}^\infty$ converges pointwise to a mapping f from X into Y .

One can verify that

$$(5.3.4) \quad \{f_j\}_{j=1}^\infty \text{ converges uniformly to } f \text{ on } X$$

in this situation. This is analogous to the fact that the space $\mathcal{B}(X, Y)$ of bounded mappings from X into Y is complete with respect to the supremum metric when $X \neq \emptyset$ and Y is complete, as in Section 1.12. More precisely,

$$(5.3.5) \quad \begin{aligned} d_Y(f_j(x), f(x)) &\leq d_Y(f_j(x), f_l(x)) + d_Y(f_l(x), f(x)) \\ &< \epsilon + d_Y(f_l(x), f(x)) \end{aligned}$$

for all $x \in X$ and $j, l \geq L(\epsilon)$, using the triangle inequality in the first step, and (5.3.1) in the second step. This implies that

$$(5.3.6) \quad d_Y(f_j(x), f(x)) \leq \epsilon$$

for all $x \in X$ and $j \geq L(\epsilon)$, because $\{f_l(x)\}_{l=1}^\infty$ converges to $f(x)$ in Y .

5.3.3 Equicontinuity of uniformly Cauchy sequences

Suppose now that (X, d_X) is a metric space too, and let $\{f_j\}_{j=1}^\infty$ be a uniformly Cauchy sequence of mappings from X into Y . Let $x \in X$ be given, and suppose that f_j is continuous at x for each $j \in \mathbf{Z}_+$. Under these conditions,

$$(5.3.7) \quad \{f_j\}_{j=1}^\infty \text{ is equicontinuous at } x.$$

To see this, let $\epsilon > 0$ be given, and let $L(\epsilon/3) \in \mathbf{Z}_+$ be as before, so that

$$(5.3.8) \quad d_Y(f_j(w), f_l(w)) < \epsilon/3$$

for every $w \in X$ and $j, l \geq L(\epsilon/3)$. It follows that

$$\begin{aligned} d_Y(f_j(x), f_j(w)) &\leq d_Y(f_j(x), f_l(x)) + d_Y(f_l(x), f_l(w)) + d_Y(f_l(w), f_j(w)) \\ (5.3.9) \quad &< d_Y(f_l(x), f_l(w)) + 2\epsilon/3 \end{aligned}$$

for every $w \in X$ and $j, l \geq L(\epsilon/3)$, using the triangle inequality twice in the first step.

If l is any positive integer, then we can use the continuity of f_l at x to get a $\delta_l(x, \epsilon/3) > 0$ such that

$$(5.3.10) \quad d_Y(f_l(x), f_l(w)) < \epsilon/3$$

for every $w \in X$ with

$$(5.3.11) \quad d_X(x, w) < \delta_l(x, \epsilon/3).$$

If $l \geq L(\epsilon/3)$, then we can combine this with (5.3.9) to get that

$$(5.3.12) \quad d_Y(f_j(x), f_j(w)) < \epsilon$$

for every $j \geq L(\epsilon/3)$ and $w \in X$ such that (5.3.11) holds. We may as well take $l = L(\epsilon/3)$ here.

We would like to find a $\delta(x, \epsilon) > 0$ such that (5.3.12) holds for every $j \in \mathbf{Z}_+$ and $w \in X$ with

$$(5.3.13) \quad d_X(x, w) < \delta(x, \epsilon).$$

This can be obtained from the previous statement for $j \geq L(\epsilon/3)$ and the continuity of f_j at x when $j < L(\epsilon/3)$.

Similarly, if f_j is uniformly continuous on X for each $j \geq 1$, then

$$(5.3.14) \quad \{f_j\}_{j=1}^{\infty} \text{ is uniformly equicontinuous on } X.$$

5.4 Equicontinuity and uniform convergence

Let (X, d_X) and (Y, d_Y) be metric spaces again, and let $\{f_j\}_{j=1}^{\infty}$ be a sequence of mappings from X into Y . Suppose that

$$(5.4.1) \quad \{f_j\}_{j=1}^{\infty} \text{ converges pointwise to a mapping } f \text{ from } X \text{ into } Y,$$

and that

$$(5.4.2) \quad \{f_j\}_{j=1}^{\infty} \text{ is equicontinuous at every } x \in X.$$

5.4.1 A localized uniform convergence property

Let $\epsilon > 0$ be given, so that for each $x \in X$ there is a $\delta(x, \epsilon) > 0$ such that

$$(5.4.3) \quad d_Y(f_j(x), f_j(w)) < \epsilon/3$$

for every $j \in \mathbf{Z}_+$ and $w \in X$ with $d_X(x, w) < \delta(x, \epsilon)$. This implies that

$$(5.4.4) \quad d_Y(f(x), f(w)) \leq \epsilon/3$$

for every $x, w \in X$ with $d_X(x, w) < \delta(x, \epsilon)$, as in Subsection 5.2.2.

It follows that

$$\begin{aligned}
 (5.4.5) \quad d_Y(f_j(w), f(w)) & \\
 & \leq d_Y(f_j(w), f_j(x)) + d_Y(f_j(x), f(x)) + d_Y(f(x), f(w)) \\
 & < 2\epsilon/3 + d_Y(f_j(x), f(x))
 \end{aligned}$$

for every $j \in \mathbf{Z}_+$ and $x, w \in X$ with $d_X(x, w) < \delta(x, \epsilon)$, using the triangle inequality twice in the first step, and (5.4.3), (5.4.4) in the second step.

If $x \in X$, then there is a positive integer $L(x, \epsilon)$ such that

$$(5.4.6) \quad d_Y(f_j(x), f(x)) < \epsilon/3$$

for every $j \geq L(x, \epsilon)$, because of (5.4.1). Combining this with (5.4.5), we obtain that

$$(5.4.7) \quad d_Y(f_j(w), f(w)) < \epsilon$$

for every $x \in X$, $j \geq L(x, \epsilon)$, and $w \in X$ with $d_X(x, w) < \delta(x, \epsilon)$.

5.4.2 Uniform convergence on compact sets

Let K be a compact subset of X . The collection of open balls

$$(5.4.8) \quad B_X(x, \delta(x, \epsilon))$$

in X centered at elements x of K is an open covering of K in X . Because K is compact, there are finitely many elements x_1, \dots, x_n of K such that

$$(5.4.9) \quad K \subseteq \bigcup_{l=1}^n B_X(x_l, \delta(x_l, \epsilon)).$$

Put

$$(5.4.10) \quad L_K(\epsilon) = \max_{1 \leq l \leq n} L(x_l, \epsilon).$$

It follows that (5.4.7) holds for every $w \in K$ and $j \geq L_K(\epsilon)$. This means that

$$(5.4.11) \quad \{f_j\}_{j=1}^\infty \text{ converges to } f \text{ uniformly on } K$$

under these conditions.

5.4.3 Another uniform convergence property

If $\{f_j\}_{j=1}^\infty$ is uniformly equicontinuous on X , then we can take $\delta(x, \epsilon) = \delta(\epsilon)$ to be independent of $x \in X$. In this case, one can check that

$$(5.4.12) \quad \{f_j\}_{j=1}^\infty \text{ converges to } f \text{ uniformly on totally bounded subsets of } X,$$

using an argument like the one in the preceding subsection.

5.5 Equicontinuity and Cauchy sequences

Let (X, d_X) , (Y, d_Y) be metric spaces, and let $\{f_j\}_{j=1}^\infty$ be a sequence of mappings from X into Y . Also let E be a subset of X , and suppose that for every $w \in E$,

$$(5.5.1) \quad \{f_j(w)\}_{j=1}^\infty \text{ is a Cauchy sequence in } Y.$$

If $x \in X$ is a limit point of E , and if

$$(5.5.2) \quad \{f_j\}_{j=1}^\infty \text{ is equicontinuous at } x,$$

then

$$(5.5.3) \quad \{f_j(x)\}_{j=1}^\infty \text{ is a Cauchy sequence in } Y.$$

If Y is complete as a metric space, then it follows that $\{f_j(x)\}_{j=1}^\infty$ converges in Y .

5.5.1 Some initial steps

To see this, let $\epsilon > 0$ be given. Because of (5.5.2), there is a $\delta > 0$ such that

$$(5.5.4) \quad d_Y(f_j(x), f_j(w)) < \epsilon/3$$

for every $j \geq 1$ and $w \in X$ with $d_X(x, w) < \delta$. Thus

$$\begin{aligned} (5.5.5) \quad d_Y(f_j(x), f_l(x)) & \leq d_Y(f_j(x), f_j(w)) + d_Y(f_j(w), f_l(w)) + d_Y(f_l(w), f_l(x)) \\ & < 2\epsilon/3 + d_Y(f_j(w), f_l(w)) \end{aligned}$$

for every $j, l \geq 1$ and $w \in X$ with $d_X(x, w) < \delta$. This uses the triangle inequality twice in the first step, and (5.5.4) twice in the second step, applied to j and l .

5.5.2 Using the Cauchy condition at a nearby point

If x is a limit point of E , then there is a $w \in E$ such that $d_X(x, w) < \delta$. In this case, (5.5.1) implies that there is a positive integer L such that

$$(5.5.6) \quad d_Y(f_j(w), f_l(w)) < \epsilon/3$$

for every $j, l \geq L$. Combining this with (5.5.5), we get that

$$(5.5.7) \quad d_Y(f_j(x), f_l(x)) < \epsilon$$

for every $j, l \geq L$, as desired.

5.6 Equicontinuity and subsequences

Let (X, d_X) and (Y, d_Y) be metric spaces again, and let $\{f_j\}_{j=1}^\infty$ be a sequence of mappings from X into Y . Suppose that for every $x \in X$ there is a sequentially compact set $K(x) \subseteq Y$ such that

$$(5.6.1) \quad f_j(x) \in K(x)$$

for every $j \geq 1$. Suppose also that

$$(5.6.2) \quad X \text{ is separable}$$

as a metric space, and let E be a dense set in X with only finitely or countably many elements. Thus

$$(5.6.3) \quad \begin{array}{l} \text{there is a subsequence } \{f_{j_l}\}_{l=1}^\infty \text{ of } \{f_j\}_{j=1}^\infty \\ \text{that converges pointwise on } E, \end{array}$$

as in Section 5.1.

5.6.1 Pointwise convergence on X

Of course, if

$$(5.6.4) \quad \{f_j\}_{j=1}^\infty \text{ is equicontinuous at every } x \in X,$$

then

$$(5.6.5) \quad \{f_{j_l}\}_{l=1}^\infty \text{ is equicontinuous at every } x \in X$$

too. In this case,

$$(5.6.6) \quad \{f_{j_l}(x)\}_{l=1}^\infty \text{ is a Cauchy sequence in } Y$$

for every $x \in X$, because E is dense in X , as in the previous section. It follows that

$$(5.6.7) \quad \{f_{j_l}(x)\}_{l=1}^\infty \text{ converges to an element of } K(x)$$

for each $x \in X$, because a Cauchy sequence of elements of a sequentially compact set converges to an element of that set, as in Section 4.10.

5.6.2 The limit f of $\{f_{j_l}\}_{l=1}^\infty$

Let $f(x)$ be the limit of $\{f_{j_l}(x)\}_{l=1}^\infty$ for every $x \in X$. Thus f is a mapping from X into Y , and

$$(5.6.8) \quad \{f_{j_l}\}_{l=1}^\infty \text{ converges to } f \text{ pointwise on } X.$$

Note that

$$(5.6.9) \quad f \text{ is continuous on } X,$$

because of (5.6.5), as in Subsection 5.2.2. We also get that

$$(5.6.10) \quad \{f_{j_l}\}_{l=1}^\infty \text{ converges to } f \text{ uniformly on compact subsets of } X,$$

as in Subsection 5.4.2.

Similarly, if

$$(5.6.11) \quad \{f_j\}_{j=1}^{\infty} \text{ is uniformly equicontinuous on } X,$$

then

$$(5.6.12) \quad \{f_{j_l}\}_{l=1}^{\infty} \text{ is uniformly equicontinuous on } X$$

as well. Under these conditions,

$$(5.6.13) \quad f \text{ is uniformly continuous on } X,$$

as in Subsection 5.2.3. In addition,

$$(5.6.14) \quad \{f_{j_l}\}_{l=1}^{\infty} \text{ converges to } f \text{ uniformly on totally bounded subsets of } X,$$

as in Subsection 5.4.3.

Of course, this is all a bit simpler when X is compact. Note that X is automatically separable in this case, because X is totally bounded, as in Section 4.2 and Subsection 4.3.1.

5.6.3 Some additional remarks

Let us now take $X = [0, 1]$ and $Y = \mathbf{R}$, using the standard Euclidean metric on \mathbf{R} and its restriction to $[0, 1]$. Put $f_j(x) = x^j$ on $[0, 1]$ for every $j \geq 1$, and remember that $\{f_j\}_{j=1}^{\infty}$ converges pointwise on $[0, 1]$, as in Subsection 1.8.1. One can check that $\{f_j\}_{j=1}^{\infty}$ is equicontinuous at $x \in [0, 1]$ when $x < 1$, and not when $x = 1$. More precisely, if $0 \leq r < 1$, then $\{f_j\}_{j=1}^{\infty}$ is uniformly equicontinuous on $[0, r]$, and $\{f_j\}_{j=1}^{\infty}$ converges to 0 uniformly on $[0, r]$. However, there is no subsequence of $\{f_j\}_{j=1}^{\infty}$ that converges uniformly on $[0, 1]$.

Another example like this is mentioned at the beginning of Section A.4. In this example, one has a uniformly bounded sequence of continuous real-valued functions on $[0, 1]$ that converges to 0 pointwise, and no subsequence converges uniformly on $[0, 1]$.

One can use uniformly convergent subsequences to get existence of solutions to ordinary differential equations under suitable conditions, as in a famous theorem of Peano. This is related to Exercises 25, 26 on p170f of [210], for instance. Note that uniqueness of solutions involves additional conditions on the differential equation, as in Exercises 27, 28 on p119 of [210].

The question of finding more elementary approaches to existence theorems like this was raised in [134]. Some responses to this question can be found in [82, 250, 251].

5.7 Pointwise and uniform boundedness

Let X be a set, let (Y, d_Y) be a metric space, and let \mathcal{E} be a collection of mappings from X into Y . If $x \in X$, then put

$$(5.7.1) \quad \mathcal{E}(x) = \{f(x) : f \in \mathcal{E}\},$$

which is a subset of Y . Let us say that \mathcal{E} is *pointwise bounded* on a subset A of X if for every $x \in A$,

$$(5.7.2) \quad \mathcal{E}(x) \text{ is bounded in } Y.$$

Similarly, put

$$(5.7.3) \quad \mathcal{E}(A) = \bigcup_{x \in A} \mathcal{E}(x) = \bigcup_{f \in \mathcal{E}} f(A) = \{f(x) : x \in A, f \in \mathcal{E}\},$$

which is a subset of Y as well. We say that \mathcal{E} is *uniformly bounded* on A when

$$(5.7.4) \quad \mathcal{E}(A) \text{ is bounded in } Y.$$

If \mathcal{E} is uniformly bounded on A , then \mathcal{E} is pointwise bounded on A , and the restriction of every $f \in \mathcal{E}$ to A is bounded as a mapping from A into Y .

Suppose that $X, Y \neq \emptyset$, and let \mathcal{E} be a collection of bounded mappings from X into Y . One can check that

$$(5.7.5) \quad \mathcal{E} \text{ is uniformly bounded on } X$$

if and only if \mathcal{E} is bounded as a subset of the space $\mathcal{B}(X, Y)$ of bounded mappings from X into Y , with respect to the supremum metric.

5.7.1 Pointwise boundedness and equicontinuity

Now let (X, d_X) be a metric space too, and let \mathcal{E} be a collection of mappings from X into Y that is equicontinuous at a point $x \in X$. Let $\epsilon > 0$ be given, so that there is a $\delta(x, \epsilon) > 0$ such that

$$(5.7.6) \quad d_Y(f(x), f(w)) < \epsilon$$

for every $f \in \mathcal{E}$ and $w \in X$ with $d_X(x, w) < \delta(x, \epsilon)$. If there is a $w_0 \in X$ such that $d_X(x, w_0) < \delta(x, \epsilon)$ and

$$(5.7.7) \quad \mathcal{E}(w_0) \text{ is bounded in } Y,$$

then it is easy to see that (5.7.2) holds.

Similarly, if (5.7.2) holds, then

$$(5.7.8) \quad \mathcal{E} \text{ is uniformly bounded on the open ball } B_X(x, \delta(x, \epsilon)).$$

5.7.2 Equicontinuity on X

Suppose for the moment that

$$(5.7.9) \quad \mathcal{E} \text{ is equicontinuous at every } x \in X.$$

One can check that

$$(5.7.10) \quad \{u \in X : \mathcal{E}(u) \text{ is bounded in } Y\} \text{ is a closed set in } X,$$

using the remarks in the preceding subsection. More precisely, every limit point of this set in X is an element of this set, by the previous remarks.

If $K \subseteq X$ is compact, and

$$(5.7.11) \quad \mathcal{E} \text{ is pointwise bounded on } K,$$

then one can verify that

$$(5.7.12) \quad \mathcal{E} \text{ is uniformly bounded on } K,$$

using (5.7.9). Indeed, (5.7.8) implies that K can be covered by open balls on which \mathcal{E} is uniformly bounded. If K is compact, then K can be covered by finitely many of these open balls. One can use this to get that $\mathcal{E}(K)$ is a bounded set in Y , because the union of finitely many bounded subsets of Y is a bounded set too, as in Subsection 1.9.2.

If

$$(5.7.13) \quad \mathcal{E} \text{ is uniformly equicontinuous on } X,$$

then we can take $\delta(x, \epsilon) = \delta(\epsilon)$ to be independent of $x \in X$ in the previous subsection. In this case, if

$$(5.7.14) \quad A \subseteq X \text{ is totally bounded}$$

and

$$(5.7.15) \quad \mathcal{E} \text{ is pointwise bounded on } A,$$

then

$$(5.7.16) \quad \mathcal{E} \text{ is uniformly bounded on } A.$$

In fact, one can cover A by finitely many open balls of the same radius on which \mathcal{E} is uniformly bounded under these conditions.

Suppose that (5.7.9) holds again. One can check that the set of $u \in X$ such that $\mathcal{E}(u)$ is totally bounded in Y is a closed set in X . This uses the equicontinuity condition for all $\epsilon > 0$, while a single $\epsilon > 0$ would suffice for the previous remarks about boundedness. If $K \subseteq X$ is compact, and $\mathcal{E}(x)$ is totally bounded in Y for every $x \in K$, then $\mathcal{E}(K)$ is totally bounded in Y . If \mathcal{E} is uniformly equicontinuous on X , $A \subseteq X$ is totally bounded, and $\mathcal{E}(x)$ is totally bounded in Y for every $x \in A$, then $\mathcal{E}(A)$ is totally bounded in Y .

5.8 Total boundedness in $\mathcal{B}(X, Y)$

Let X be a nonempty set, and let (Y, d_Y) be a nonempty metric space. Remember that $\mathcal{B}(X, Y)$ is the space of bounded mappings from X into Y , and that $\theta(f, g)$ denotes the supremum metric on $\mathcal{B}(X, Y)$, as in Subsection 1.11.2. Suppose that

$$(5.8.1) \quad \mathcal{E} \subseteq \mathcal{B}(X, Y) \text{ is totally bounded with respect to } \theta(\cdot, \cdot).$$

Suppose that (X, d_X) is a metric space too, and that

$$(5.8.2) \quad \text{every } f \in \mathcal{E} \text{ is continuous at } x.$$

We would like to check that

$$(5.8.3) \quad \mathcal{E} \text{ is equicontinuous at } x.$$

5.8.1 Equicontinuity using total boundedness

Let $\epsilon > 0$ be given. Because \mathcal{E} is totally bounded with respect to the supremum metric, there are finitely many elements f_1, \dots, f_n of \mathcal{E} such that for every $f \in \mathcal{E}$ there is a positive integer $j \leq n$ such that

$$(5.8.4) \quad \theta(f, f_j) \leq \epsilon/3.$$

By hypothesis, f_j is continuous at x for each $j = 1, \dots, n$, so that there is a $\delta_j(x) > 0$ such that

$$(5.8.5) \quad d_Y(f_j(x), f_j(w)) < \epsilon/3$$

for every $w \in X$ with $d_X(x, w) < \delta_j(x)$. Put

$$(5.8.6) \quad \delta(x) = \min_{1 \leq j \leq n} \delta_j(x) > 0.$$

Let $f \in \mathcal{E}$ be given, and let $j \leq n$ be as in (5.8.4). Thus

$$(5.8.7) \quad \begin{aligned} d_Y(f(x), f(w)) &\leq d_Y(f(x), f_j(x)) + d_Y(f_j(x), f_j(w)) + d_Y(f_j(w), f(w)) \\ &\leq d_Y(f_j(x), f_j(w)) + 2\epsilon/3 \end{aligned}$$

for every $w \in X$, using the triangle inequality in the first step.

It follows that

$$(5.8.8) \quad d_Y(f(x), f(w)) < \epsilon/3 + 2\epsilon/3 = \epsilon$$

when $d_X(x, w) < \delta(x) \leq \delta_j(x)$, using (5.8.5) in the first step. This shows that (5.8.3) holds, as desired.

Similarly, if

$$(5.8.9) \quad \text{every } f \in \mathcal{E} \text{ is uniformly continuous on } X,$$

then

$$(5.8.10) \quad \mathcal{E} \text{ is uniformly equicontinuous on } X.$$

This is essentially the same as before, but with $\delta_j(x)$ and hence $\delta(x)$ independent of x .

5.8.2 Some more total boundedness conditions

Let us return to the hypotheses at the beginning of the section, so that (5.8.1) holds, without any additional equicontinuity conditions. Let $x \in X$ be given, and remember that $\mathcal{E}(x)$ is the subset of Y defined in (5.7.1). It is easy to see that

$$(5.8.11) \quad \mathcal{E}(x) \text{ is totally bounded in } Y.$$

Let A be a subset of X , and remember that $\mathcal{E}(A)$ is the subset of Y defined in (5.7.3). Suppose that

$$(5.8.12) \quad \text{for each } f \in \mathcal{E}, f(A) \text{ is totally bounded in } Y.$$

One can check that

$$(5.8.13) \quad \mathcal{E}(A) \text{ is totally bounded in } Y$$

as well under these conditions. Of course, (5.8.1) holds automatically when \mathcal{E} has only finitely many elements.

5.9 Total boundedness using equicontinuity

Let (X, d_X) and (Y, d_Y) be nonempty metric spaces again. If f is a uniformly continuous mapping from X into Y and A is a totally bounded subset of X , then it is not difficult to show that

$$(5.9.1) \quad f(A) \text{ is totally bounded in } Y.$$

Suppose for the rest of the section that

$$(5.9.2) \quad X \text{ is totally bounded with respect to } d_X.$$

This implies that uniformly continuous mappings from X into Y are bounded, because totally bounded subsets of Y are bounded. Let $\theta(\cdot, \cdot)$ be the supremum metric on the space $\mathcal{B}(X, Y)$ of bounded mappings from X into Y , as usual.

5.9.1 Using uniform equicontinuity

Suppose that

$$(5.9.3) \quad \mathcal{E} \text{ is a uniformly equicontinuous collection of mappings from } X \text{ into } Y.$$

In particular, every $f \in \mathcal{E}$ is uniformly continuous as a mapping from X into Y , and hence bounded, as in the preceding paragraph.

Let $\epsilon > 0$ be given, so that there is a $\delta > 0$ such that

$$(5.9.4) \quad d_Y(f(x), f(w)) < \epsilon/3$$

holds for every $f \in \mathcal{E}$ and $x, w \in X$ with $d_X(x, w) < \delta$. Because X is totally bounded, there are finitely many elements x_1, \dots, x_n of X such that

$$(5.9.5) \quad X = \bigcup_{j=1}^n B_X(x_j, \delta),$$

where $B_X(x, r)$ is the open ball in X centered at $x \in X$ with radius $r > 0$, as usual.

Let $f, g \in \mathcal{E}$ and $x \in X$ be given, so that $d_X(x_j, x) < \delta$ for some $j \leq n$, by (5.9.5). Under these conditions, we have that

$$(5.9.6) \quad d_Y(f(x_j), f(x)), d_Y(g(x_j), g(x)) < \epsilon/3,$$

as in (5.9.4).

This implies that

$$(5.9.7) \quad \begin{aligned} d_Y(f(x), g(x)) &\leq d_Y(f(x), f(x_j)) + d_Y(f(x_j), g(x_j)) + d_Y(g(x_j), g(x)) \\ &< d_Y(f(x_j), g(x_j)) + 2\epsilon/3, \end{aligned}$$

using the triangle inequality in the first step. Thus

$$(5.9.8) \quad d_Y(f(x), g(x)) < \max_{1 \leq j \leq n} d_Y(f(x_j), g(x_j)) + 2\epsilon/3$$

for every $x \in X$. It follows that

$$(5.9.9) \quad \theta(f, g) \leq \max_{1 \leq j \leq n} d_Y(f(x_j), g(x_j)) + 2\epsilon/3$$

for every $f, g \in \mathcal{E}$.

5.9.2 Total boundedness of \mathcal{E}

Suppose that

$$(5.9.10) \quad \mathcal{E}(x) = \{f(x) : f \in \mathcal{E}\}$$

is totally bounded in Y for each $x \in X$, in addition to (5.9.2) and (5.9.3). Using this and (5.9.9), one can check that

$$(5.9.11) \quad \mathcal{E} \text{ is totally bounded with respect to the supremum metric.}$$

This corresponds to Theorem 10.4 C on p269 of [93] when X is a closed interval in the real line, and $Y = \mathbf{R}$. Another version of this is given in Theorem A 5 on p369 of [208], with X a compact topological space, and $Y = \mathbf{C}$.

5.10 Equiconvergence of limits

Let (X, d_X) and (Y, d_Y) be metric spaces, let A be a subset of X , and suppose that $x \in X$ is a limit point of A in X . Also let \mathcal{E} be a collection of mappings from E into Y , and suppose that for each $f \in \mathcal{E}$, the limit

$$(5.10.1) \quad \lim_{\substack{w \in A \\ w \rightarrow x}} f(w) = q_f$$

exists in Y . Let us say that we have *equiconvergence of the limit* for $f \in \mathcal{E}$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$(5.10.2) \quad d_Y(f(w), q_f) < \epsilon$$

for every $f \in \mathcal{E}$ and $w \in A$ with $d_X(x, w) < \delta$ and $w \neq x$.

More precisely, this condition includes the existence of the limit as in (5.10.1) for each $f \in \mathcal{E}$. If \mathcal{E} has only finitely many elements, and the limit as in (5.10.1) exists for each $f \in \mathcal{E}$, then one can check that we have equiconvergence of the limit for $f \in \mathcal{E}$.

5.10.1 Equiconvergence and equicontinuity

Suppose for the moment that $A = X$, so that x is a limit point of X . In this case, it is well known and easy to see that a mapping f from X into Y is continuous at x if and only if

$$(5.10.3) \quad \lim_{w \rightarrow x} f(w) = f(x),$$

where the existence of the limit is part of this condition. One can verify that we have equiconvergence of the limit for $f \in \mathcal{E}$, with $q_f = f(x)$ for every $f \in \mathcal{E}$, exactly when \mathcal{E} is equicontinuous at x .

5.10.2 Equiconvergence for sequences of mappings

Let A be any subset of X again, with $x \in X$ a limit point of A , and let $\{f_j\}_{j=1}^{\infty}$ be a sequence of mappings from A into Y . Suppose that for each positive integer j , the limit

$$(5.10.4) \quad \lim_{\substack{w \in A \\ w \rightarrow x}} f_j(w) = q_j$$

exists in Y . We can define *equiconvergence of the limit* for $\{f_j\}_{j=1}^{\infty}$ in the same way as before, by considering the collection of f_j 's, $j \in \mathbf{Z}$.

Suppose for the moment that

$$(5.10.5) \quad \text{for each } w \in A, \{f_j(w)\}_{j=1}^{\infty} \text{ is a Cauchy sequence in } Y.$$

If we have equiconvergence of the limit in (5.10.4) for $\{f_j\}_{j=1}^{\infty}$, then one can show that

$$(5.10.6) \quad \{q_j\}_{j=1}^{\infty} \text{ is a Cauchy sequence in } Y.$$

This is very similar to the argument in Section 5.5. If Y is complete as a metric space, then it follows that

$$(5.10.7) \quad \{q_j\}_{j=1}^{\infty} \text{ converges to an element } q \text{ of } Y.$$

5.10.3 Pointwise convergent sequences

Suppose now that $\{f_j\}_{j=1}^\infty$ converges pointwise to a mapping f from A into Y . If we have equiconvergence of the limit in (5.10.4) for $\{f_j\}_{j=1}^\infty$, and if (5.10.7) holds, then one can show that

$$(5.10.8) \quad \lim_{\substack{w \in A \\ w \rightarrow x}} f(w) = q.$$

This is analogous to the argument for equicontinuous sequences of functions that converge pointwise, as in Subsection 5.2.2.

5.10.4 Equiconvergence and uniform Cauchy sequences

Let $\{f_j\}_{j=1}^\infty$ be a uniformly Cauchy sequence of mappings from A into Y , as in Section 5.3. Suppose that the limit as in (5.10.4) exists for each $j \in \mathbf{Z}_+$. Under these conditions, one can check that (5.10.6) holds. This was mentioned in Subsection 1.8.4 for uniformly convergent sequences of mappings.

Using this, one can verify that we have equiconvergence of the limit as in (5.10.4) for $\{f_j\}_{j=1}^\infty$. This is analogous to the argument for equicontinuity in Subsection 5.3.3.

Suppose that $\{f_j\}_{j=1}^\infty$ converges uniformly to a mapping f from A into Y , so that $\{f_j\}_{j=1}^\infty$ is uniformly Cauchy on A in particular, as in Subsection 5.3.1. Suppose also that the limit as in (5.10.4) exists for every j , and that (5.10.7) holds. Using the remarks in the previous two paragraphs and the previous subsection, we get that (5.10.8) holds too. This was mentioned in Subsection 1.8.4 as well.

5.11 Equiconvergence and differentiability

Let a, b be real numbers with $a < b$, and let f be a real-valued function on $[a, b]$. As usual, the derivative of f at $x \in [a, b]$ is defined by

$$(5.11.1) \quad f'(x) = \lim_{w \rightarrow x} \frac{f(w) - f(x)}{w - x},$$

when the limit exists. Of course, this is really a one-sided limit and derivative when $x = a$ or b .

Let \mathcal{E} be a collection of real-valued functions on $[a, b]$, each of which is differentiable at $x \in [a, b]$. In this case, we may be interested in the equiconvergence of the limit of difference quotients in (5.11.1) for $f \in \mathcal{E}$, as in the previous section.

5.11.1 A criterion for equiconvergence

Suppose for the moment that each $f \in \mathcal{E}$ is differentiable at every point in $[a, b]$. If the collection

$$(5.11.2) \quad \mathcal{E}' = \{f' : f \in \mathcal{E}\}$$

of derivatives of elements of \mathcal{E} is equicontinuous at x , then one can check that we have equiconvergence of the limit as in (5.11.1) for $f \in \mathcal{E}$, using the mean value theorem.

5.11.2 Sequences of differentiable mappings

Let $\{f_j\}_{j=1}^\infty$ be a sequence of real-valued functions on $[a, b]$, and suppose that f_j is differentiable at a point $x \in [a, b]$ for each j . As before, we may be interested in the equiconvergence of the limit

$$(5.11.3) \quad f'_j(x) = \lim_{w \rightarrow x} \frac{f_j(w) - f_j(x)}{w - x}$$

for $\{f_j\}_{j=1}^\infty$.

Suppose for the moment that f_j is differentiable at every point in $[a, b]$ for each j . If $\{f_j\}_{j=1}^\infty$ is equicontinuous at x , then we have equiconvergence of the limit as in (5.11.3) for $\{f_j\}_{j=1}^\infty$, as before.

5.11.3 Pointwise convergence and differentiability

Suppose now that $\{f_j\}_{j=1}^\infty$ converges pointwise to a real-valued function f on $[a, b]$. If $x, w \in [a, b]$ and $x \neq w$, then it follows that

$$(5.11.4) \quad \lim_{j \rightarrow \infty} \frac{f_j(w) - f_j(x)}{w - x} = \frac{f(w) - f(x)}{w - x}.$$

If f_j is differentiable at x for each j , and we have equiconvergence of the limit as in (5.11.3) for $\{f_j\}_{j=1}^\infty$, then $\{f'_j(x)\}_{j=1}^\infty$ is a Cauchy sequence in \mathbf{R} , as in Subsection 5.10.2. Of course, this means that $\{f'_j(x)\}_{j=1}^\infty$ converges in \mathbf{R} , because \mathbf{R} is complete with respect to the standard Euclidean metric.

Using this, we get that f is differentiable at x , with

$$(5.11.5) \quad f'(x) = \lim_{j \rightarrow \infty} f'_j(x),$$

as in Subsection 5.10.3.

5.11.4 Uniformly Cauchy difference quotients

Suppose that f_j is differentiable at a point $x \in [a, b]$ for each j again. Suppose also that the sequence of difference quotients

$$(5.11.6) \quad \frac{f_j(w) - f_j(x)}{w - x}$$

is uniformly Cauchy as a sequence of real-valued functions of w on $[a, b] \setminus \{x\}$. This implies that $\{f'_j(x)\}_{j=1}^\infty$ is a Cauchy sequence, as in Subsection 5.10.4.

This means that $\{f'_j(x)\}_{j=1}^\infty$ converges in \mathbf{R} , because \mathbf{R} is complete with respect to the standard Euclidean metric. In this case, we get equiconvergence of the limit as in (5.11.3) for $\{f_j\}_{j=1}^\infty$ too, as in Subsection 5.10.4 again.

Note that

$$\begin{aligned}
 (5.11.7) \quad & \frac{f_j(w) - f_j(x)}{w - x} - \frac{f_l(w) - f_l(x)}{w - x} \\
 &= \frac{(f_j(w) - f_l(w)) - (f_j(x) - f_l(x))}{w - x}
 \end{aligned}$$

for every $w \in [a, b] \setminus \{x\}$ and $j, l \geq 1$. The uniform Cauchy condition for (5.11.6) on $[a, b] \setminus \{x\}$ means that if j, l are sufficiently large, then (5.11.7) is as small as we like, uniformly over $w \in [a, b] \setminus \{x\}$.

Suppose that f_j is differentiable at every point in $[a, b]$ for each j , and that $\{f'_j\}_{j=1}^\infty$ is uniformly Cauchy as a sequence of real-valued functions on $[a, b]$. In this case, one can use the mean value theorem to get that (5.11.7) is as small as we like when j, l are sufficiently large, uniformly over $x, w \in [a, b]$ with $x \neq w$. This is related to Theorem 7.17 on p152 of [210].

A simpler argument for continuously-differentiable functions on $[a, b]$ will be discussed in Subsection 6.3.1. This corresponds to the remark on p153 of [210].

Chapter 6

More on sums and norms

6.1 Weierstrass' criterion

Let X be a nonempty set, and let $a_1(x), a_2(x), a_3(x), \dots$ be an infinite sequence of complex-valued functions on X . Also let A_1, A_2, A_3, \dots be an infinite sequence of nonnegative real numbers such that

$$(6.1.1) \quad |a_j(x)| \leq A_j$$

for every $x \in X$ and $j \geq 1$. Suppose that

$$(6.1.2) \quad \sum_{j=1}^{\infty} A_j$$

converges. This implies that

$$(6.1.3) \quad \sum_{j=1}^{\infty} a_j(x)$$

converges absolutely for every $x \in X$, by the comparison test.

Under these conditions, the sequence of partial sums

$$(6.1.4) \quad \sum_{j=1}^n a_j(x)$$

converges to (6.1.3) uniformly on X . This is a well-known criterion of Weierstrass for uniform convergence.

To see this, observe that

$$(6.1.5) \quad \left| \sum_{j=1}^{\infty} a_j(x) - \sum_{j=1}^n a_j(x) \right| = \left| \sum_{j=n+1}^{\infty} a_j(x) \right| \leq \sum_{j=n+1}^{\infty} |a_j(x)| \leq \sum_{j=n+1}^{\infty} A_j$$

for every $x \in X$ and $n \geq 1$. This uses the triangle inequality for finite sums in the second step, and some additional arguments to extend this to absolutely

convergent series. The third step uses (6.1.1), and some additional arguments for infinite series.

The convergence of (6.1.2) implies that the right side of (6.1.5) tends to 0 as $n \rightarrow \infty$. It follows that (6.1.4) converges to (6.1.3) uniformly on X , because the right side of (6.1.5) does not depend on x .

6.1.1 Continuity of the sum

Suppose that $(X, d(\cdot, \cdot))$ is a metric space, and that $a_j(x)$ is a continuous complex-valued function on X for each $j \geq 1$, with respect to the standard Euclidean metric on the complex plane \mathbf{C} . This implies that the partial sums (6.1.4) are continuous on X for each n as well.

If (6.1.2) converges, then the partial sums converge to (6.1.3) uniformly on X , as before. It follows that (6.1.3) is continuous on X too, as in Subsection 1.8.2.

6.1.2 Power series

Now let

$$(6.1.6) \quad \sum_{j=0}^{\infty} a_j z^j$$

be a power series with complex coefficients. Suppose that

$$(6.1.7) \quad \sum_{j=0}^{\infty} |a_j| r^j$$

converges for some nonnegative real number r . This implies that (6.1.6) converges absolutely for every $z \in \mathbf{C}$ with $|z| \leq r$, by the comparison test, as before.

Using Weierstrass' criterion, we get that the partial sums

$$(6.1.8) \quad \sum_{j=0}^n a_j z^j$$

converge to (6.1.6) uniformly on the closed disk

$$(6.1.9) \quad \{z \in \mathbf{C} : |z| \leq r\}.$$

More precisely, this uses the fact that

$$(6.1.10) \quad |a_j z^j| = |a_j| |z|^j \leq |a_j| r^j$$

on (6.1.9) for each j .

Of course, (6.1.8) is continuous as a mapping from \mathbf{C} into itself for each $n \geq 0$, using the standard metric on \mathbf{C} . In particular, the restriction of (6.1.8) to (6.1.9) is continuous with respect to the restriction of the standard metric on \mathbf{C} to (6.1.9). This implies that (6.1.6) is continuous on (6.1.9) as well, as in the previous subsection.

6.1.3 Another continuity property

Suppose that $0 < \rho \leq \infty$ has the property that (6.1.7) converges when $0 \leq r < \rho$. This implies that (6.1.6) converges absolutely for every $z \in \mathbf{C}$ with $|z| < \rho$. Under these conditions, (6.1.6) defines a continuous complex-valued function on

$$(6.1.11) \quad \{z \in \mathbf{C} : |z| < \rho\},$$

with respect to the restriction of the standard metric on \mathbf{C} to (6.1.11).

To see this, let $z_0 \in \mathbf{C}$ with $|z_0| < \rho$ be given, and let us check that (6.1.6) is continuous at z_0 . Let r be a positive real number such that

$$(6.1.12) \quad |z_0| < r < \rho.$$

The remarks in the preceding subsection imply that (6.1.6) is continuous on (6.1.9). In particular, (6.1.6) is continuous at z_0 as a complex-valued function on (6.1.9). One can use this to verify that (6.1.6) is continuous at z_0 as a complex-valued function on (6.1.11), because $|z_0| < r$.

6.2 Radius of convergence

Let

$$(6.2.1) \quad \sum_{j=0}^{\infty} a_j z^j$$

be a power series with complex coefficients, and let A be the set of nonnegative real numbers r such that

$$(6.2.2) \quad \sum_{j=0}^{\infty} |a_j| r^j$$

converges. Of course, $0 \in A$.

If $r \in A$, then

$$(6.2.3) \quad [0, r] \subseteq A,$$

by the comparison test. If t is a positive real number such that $t \notin A$, then it follows that t is an upper bound for A .

6.2.1 Defining the radius of convergence

The *radius of convergence* of (6.2.1) can be defined as a nonnegative extended real number by

$$(6.2.4) \quad R = \sup A.$$

If $0 \leq t < R$, then t is not an upper bound for A , by the definition of the supremum. This implies that

$$(6.2.5) \quad t \in A,$$

as in the preceding paragraph.

If A is the set of all nonnegative real numbers, then $R = +\infty$. Otherwise, $R < +\infty$, and A is either $[0, R]$, or the set of nonnegative real numbers strictly less than R .

If $z \in \mathbf{C}$ and $|z| < R$, then $|z| \in A$, as in (6.2.5), so that (6.2.1) converges absolutely. In fact, (6.2.1) defines a continuous function on

$$(6.2.6) \quad \{z \in \mathbf{C} : |z| < R\},$$

as in Subsection 6.1.3.

6.2.2 More on R

Let t be a positive real number such that $\{|a_j| t^j\}_{j=0}^{\infty}$ is a bounded sequence of nonnegative real numbers. This means that there is a nonnegative real number C such that

$$(6.2.7) \quad |a_j| t^j \leq C$$

for every $j \geq 0$.

If r is any nonnegative real number, then we get that

$$(6.2.8) \quad |a_j| r^j \leq C (r/t)^j$$

for every $j \geq 0$. If $r < t$, then it follows that (6.2.2) converges, by comparison with the convergent geometric series $\sum_{j=0}^{\infty} (r/t)^j$. This implies that $r \in A$, so that

$$(6.2.9) \quad t \leq R.$$

If (6.2.1) converges for some $z \in \mathbf{C}$, then

$$(6.2.10) \quad \lim_{j \rightarrow \infty} a_j z^j = 0.$$

This implies that $\{a_j z^j\}_{j=0}^{\infty}$ is a bounded sequence of complex numbers, which is the same as saying that $\{|a_j| |z|^j\}_{j=0}^{\infty}$ is a bounded sequence of nonnegative real numbers. It follows that

$$(6.2.11) \quad |z| \leq R,$$

as in (6.2.9).

6.2.3 Another convergence property

Suppose that (6.2.2) converges for some positive real number r . If r_0 is a nonnegative real number strictly less than r , then it is well known that

$$(6.2.12) \quad \lim_{j \rightarrow \infty} j (r_0/r)^j = 0,$$

because $r_0/r < 1$.

In particular, $\{j (r_0/r)^j\}_{j=0}^{\infty}$ is a bounded sequence of nonnegative real numbers, so that there is a nonnegative real number C_0 such that

$$(6.2.13) \quad j (r_0/r)^j \leq C_0$$

for every $j \geq 0$. Thus

$$(6.2.14) \quad j |a_j| r_0^j \leq C_0 |a_j| r^j$$

for every $j \geq 0$, which implies that

$$(6.2.15) \quad \sum_{j=0}^{\infty} j |a_j| r_0^j$$

converges, by the comparison test.

It follows that (6.2.15) converges when r_0 is strictly less than the radius of convergence R .

6.3 Termwise differentiation

Let a, b be real numbers with $a < b$, and let $\{f_j\}_{j=1}^{\infty}$ be a sequence of continuous real or complex-valued functions on $[a, b]$ that converges uniformly to a real or complex-valued function f on $[a, b]$, as appropriate. Thus f is also continuous on $[a, b]$, as in Subsection 1.8.2. Of course, we are implicitly using the standard Euclidean metrics on \mathbf{R} and \mathbf{C} here, and the restriction of the standard Euclidean metric on \mathbf{R} to $[a, b]$.

If $x \in [a, b]$, then put

$$(6.3.1) \quad F_j(x) = \int_a^x f_j(t) dt$$

for every $j \in \mathbf{Z}_+$, and

$$(6.3.2) \quad F(x) = \int_a^x f(t) dt,$$

using standard Riemann integrals on the right sides of (6.3.1) and (6.3.2). Observe that

$$(6.3.3) \quad \begin{aligned} |F_j(x) - F(x)| &= \left| \int_a^x (f_j(t) - f(t)) dt \right| \\ &\leq \int_a^x |f_j(t) - f(t)| dt \leq \int_a^b |f_j(t) - f(t)| dt \end{aligned}$$

for every $j \geq 1$ and $x \in [a, b]$. This implies that

$$(6.3.4) \quad \{F_j\}_{j=1}^{\infty} \text{ converges uniformly to } F \text{ on } [a, b],$$

because the right side of (6.3.3) tends to 0 as $j \rightarrow \infty$, and does not depend on x . We also have that $F'_j(x) = f_j(x)$ for every $j \geq 1$ and $x \in [a, b]$, and that $F'(x) = f(x)$ for every $x \in [a, b]$, using the appropriate one-sided derivative when $x = a$ or b .

6.3.1 Continuously-differentiable functions on $[a, b]$

Now let $\{g_j\}_{j=1}^\infty$ be a sequence of continuously-differentiable real or complex-valued functions on $[a, b]$. Thus, for each $j \in \mathbf{Z}_+$, the derivative $g'_j(x)$ of g_j exists at every $x \in [a, b]$, using the appropriate one-sided derivative when $x = a$ or b , and g'_j is continuous on $[a, b]$. It follows that

$$(6.3.5) \quad g_j(x) = g_j(a) + \int_a^x g'_j(t) dt$$

for every $j \geq 1$ and $x \in [a, b]$, by the fundamental theorem of calculus.

Suppose that $\{g_j(a)\}_{j=1}^\infty$ converges to a real or complex number $g(a)$, as appropriate, and that $\{g'_j\}_{j=1}^\infty$ converges uniformly to a real or complex-valued function f on $[a, b]$, as appropriate. Note that f is continuous on $[a, b]$, as in Subsection 1.8.2. Let g be the real or complex-valued function defined on $[a, b]$ by

$$(6.3.6) \quad g(x) = g(a) + \int_a^x f(t) dt$$

for each $x \in [a, b]$. Under these conditions,

$$(6.3.7) \quad \{g_j\}_{j=1}^\infty \text{ converges uniformly to } g \text{ on } [a, b],$$

because of (6.3.4). Of course, g is continuously differentiable on $[a, b]$, with $g' = f$ on $[a, b]$.

6.3.2 Differentiating power series

Let $\sum_{j=0}^\infty a_j x^j$ be a power series with real or complex coefficients, and suppose that

$$(6.3.8) \quad \sum_{j=0}^\infty j |a_j| r^j$$

converges for some positive real number r . In particular, this implies that

$$(6.3.9) \quad \sum_{j=0}^\infty |a_j| r^j$$

converges, and we put

$$(6.3.10) \quad f(x) = \sum_{j=0}^\infty a_j x^j$$

for every $x \in \mathbf{R}$ with $|x| \leq r$. Similarly, put

$$(6.3.11) \quad \phi(x) = \sum_{j=1}^\infty j a_j x^{j-1}$$

for every $x \in \mathbf{R}$ with $|x| \leq r$, where the series on the right converges absolutely because of the convergence of (6.3.8).

Under these conditions, the partial sums of the series on the right sides of (6.3.10) and (6.3.11) converge uniformly on $[-r, r]$, as in Subsection 6.1.2. By construction, the partial sums of the right side of (6.3.11) are the same as the first derivatives of the partial sums of the right side of (6.3.10).

Using the remarks in the previous subsection, we get that f is differentiable on $[-r, r]$, with

$$(6.3.12) \quad f'(x) = \phi(x)$$

for every $x \in [-r, r]$. This uses the appropriate one-sided derivatives when $x = \pm r$, as usual.

6.3.3 More on differentiating power series

Suppose that $0 < \rho \leq +\infty$ has the property that (6.3.9) converges when $0 \leq r < \rho$. This implies that (6.3.8) converges when $0 \leq r < \rho$ too. More precisely, one can use the convergence of $\sum_{j=0}^{\infty} |a_j| t^j$ for some $r < t < \rho$ to get that (6.3.8) converges, as in Subsection 6.2.3.

In this situation, the series on the right sides of (6.3.10) and (6.3.11) converge absolutely for every $x \in \mathbf{R}$ with $|x| < \rho$, so that $f(x)$ and $\phi(x)$ may be defined on $(-\rho, \rho)$ as before. Using the remarks in the preceding subsection, we get that f is differentiable on $(-\rho, \rho)$, with derivative given by (6.3.12).

One can repeat the process, to get that f is differentiable of all orders on $(-\rho, \rho)$, where the derivatives of f are obtained by differentiating the right side of (6.3.10) termwise. The resulting power series also converge absolutely on $(-\rho, \rho)$, for the same reasons as before.

6.4 Rearrangements

Let $\sum_{j=1}^{\infty} a_j$ be an infinite series of real or complex numbers, and let π be a one-to-one mapping from the set \mathbf{Z}_+ of positive integers onto itself. Under these conditions, the infinite series

$$(6.4.1) \quad \sum_{l=1}^{\infty} a_{\pi(l)}$$

is called a *rearrangement* of $\sum_{j=1}^{\infty} a_j$.

If $a_j = 0$ for all but finitely many positive integers j , then $a_{\pi(l)} = 0$ for all but finitely many l too, and it is easy to see that

$$(6.4.2) \quad \sum_{l=1}^{\infty} a_{\pi(l)} = \sum_{j=1}^{\infty} a_j.$$

6.4.1 Rearrangements with nonnegative terms

Suppose for the moment that a_j is a nonnegative real number for each $j \geq 1$. If $n \in \mathbf{Z}_+$, then

$$(6.4.3) \quad \sum_{l=1}^n a_{\pi(l)} \leq \sum_{j=1}^N a_j$$

for every $N \geq \max_{1 \leq l \leq n} \pi(l)$.

Similarly,

$$(6.4.4) \quad \sum_{j=1}^n a_j \leq \sum_{l=1}^N a_{\pi(l)}$$

for every $N \geq \max_{1 \leq j \leq n} \pi^{-1}(j)$. This implies that $\sum_{j=1}^{\infty} a_j$ converges if and only if $\sum_{l=1}^{\infty} a_{\pi(l)}$ converges, in which case (6.4.2) holds.

6.4.2 Rearrangements and absolute convergence

If $\sum_{j=1}^{\infty} a_j$ is an infinite series of real or complex numbers, then it follows that $\sum_{j=1}^{\infty} |a_j|$ converges if and only if $\sum_{j=1}^{\infty} |a_{\pi(j)}|$ converges. This means that $\sum_{j=1}^{\infty} a_j$ converges absolutely if and only if $\sum_{l=1}^{\infty} a_{\pi(l)}$ converges absolutely.

One can check that (6.4.2) holds in this situation as well. More precisely, if the a_j 's are real numbers, then one can reduce to the previous case by expressing $\sum_{j=1}^{\infty} a_j$ as a difference of convergent series of nonnegative real numbers. If the a_j 's are complex numbers, then one can consider their real and imaginary parts.

If $\sum_{j=1}^{\infty} a_j$ is an infinite series of real numbers that converges, and does not converge absolutely, then it is well known that there are rearrangements that do not converge, and rearrangements whose sum can be any real number. This corresponds to Theorem 3.5 D and Exercises 4, 5 on p77, 80 of [93], respectively, and Theorem 3.54 on p76 of [210].

6.5 Cauchy products

Let $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ be infinite series of complex numbers, and put

$$(6.5.1) \quad c_n = \sum_{j=0}^n a_j b_{n-j}$$

for each nonnegative integer n . The infinite series $\sum_{n=0}^{\infty} c_n$ is called the *Cauchy product* of the series $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$.

It is easy to see that

$$(6.5.2) \quad \sum_{n=0}^{\infty} c_n = \left(\sum_{j=0}^{\infty} a_j \right) \left(\sum_{l=0}^{\infty} b_l \right)$$

formally. In particular, if $a_j = 0$ for all but finitely many $j \geq 0$, and $b_l = 0$ for all but finitely many $l \geq 0$, then one can check that $c_n = 0$ for all but finitely many $n \geq 0$, and that (6.5.2) holds. One way to look at this will be mentioned in Subsection 6.5.3.

Suppose for the moment that $\sum_{j=0}^{\infty} a_j z^j$ and $\sum_{l=0}^{\infty} b_l z^l$ are power series with complex coefficients. One can check that their Cauchy product is the power series

$$(6.5.3) \quad \sum_{n=0}^{\infty} c_n z^n,$$

where c_n is as in (6.5.1) for each $n \geq 0$.

6.5.1 Nonnegative a_j, b_l

Suppose for the moment again that the a_j 's and b_l 's are nonnegative real numbers, so that the c_n 's are nonnegative real numbers too. Observe that

$$(6.5.4) \quad \sum_{n=0}^N c_n \leq \left(\sum_{j=0}^N a_j \right) \left(\sum_{l=0}^N b_l \right)$$

for every nonnegative integer N . If $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ converge, then

$$(6.5.5) \quad \sum_{n=0}^N c_n \leq \left(\sum_{j=0}^{\infty} a_j \right) \left(\sum_{l=0}^{\infty} b_l \right)$$

for every $N \geq 0$. This implies that $\sum_{n=0}^{\infty} c_n$ converges, with

$$(6.5.6) \quad \sum_{n=0}^{\infty} c_n \leq \left(\sum_{j=0}^{\infty} a_j \right) \left(\sum_{l=0}^{\infty} b_l \right).$$

We also have that

$$(6.5.7) \quad \left(\sum_{j=0}^N a_j \right) \left(\sum_{l=0}^N b_l \right) \leq \sum_{n=0}^{2N} c_n \leq \sum_{n=0}^{\infty} c_n$$

for every $N \geq 0$. If $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ converge, then we get that

$$(6.5.8) \quad \left(\sum_{j=0}^{\infty} a_j \right) \left(\sum_{l=0}^{\infty} b_l \right) \leq \sum_{n=0}^{\infty} c_n.$$

Of course, (6.5.2) follows from (6.5.6) and (6.5.8) in this situation.

6.5.2 Absolutely convergent series

Suppose now that $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ are absolutely convergent series of complex numbers. Clearly

$$(6.5.9) \quad |c_n| \leq \sum_{j=0}^n |a_j| |b_{n-j}|$$

for each $n \geq 0$, by the triangle inequality. The right side of (6.5.9) is the same as the n th term of the Cauchy product of $\sum_{j=0}^{\infty} |a_j|$ and $\sum_{l=0}^{\infty} |b_l|$. These two series converge, by hypothesis, and so their Cauchy product converges as well, as in the previous subsection.

This implies that $\sum_{n=0}^{\infty} c_n$ converges absolutely, by the comparison test. We also get that

$$(6.5.10) \quad \sum_{n=0}^{\infty} |c_n| \leq \sum_{n=0}^{\infty} \left(\sum_{j=0}^n |a_j| |b_l| \right) = \left(\sum_{j=0}^{\infty} |a_j| \right) \left(\sum_{l=0}^{\infty} |b_l| \right)$$

under these conditions.

If the a_j 's and b_l 's are real numbers, then $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ can be expressed as differences of convergent series of nonnegative real numbers. In this case, (6.5.2) can be obtained from the corresponding statement for nonnegative real numbers, as in the previous subsection.

If the a_j 's and b_l 's are complex numbers, then one can get (6.5.2) using the analogous statements for the real and imaginary parts of the a_j 's and b_l 's.

6.5.3 Another type of argument

Alternatively, consider

$$(6.5.11) \quad \sum_{j,l \geq 0} a_j b_l,$$

where more precisely the sum is taken over all ordered pairs (j, l) of nonnegative integers j, l . This sum can be identified formally with both sides of (6.5.2). The left side of (6.5.2) corresponds to summing first over (j, l) such that $j + l = n$, and then summing over $n \geq 0$. The right side of (6.5.2) can be obtained by summing over j and l separately.

If $a_j = 0$ for all but finitely many $j \geq 0$, and $b_l = 0$ for all but finitely many $l \geq 0$, then $a_j b_l = 0$ for all but finitely many (j, l) , and all of these sums can be reduced to finite sums.

If the a_j 's and b_l 's are nonnegative real numbers, then (6.5.11) can be defined as a nonnegative extended real number, as in Section 11.2. This sum can be expressed in terms of iterated sums, as in Section 11.15. In particular, one can check that (6.5.11) is finite when $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ converge, in which case (6.5.11) is the same as both sides of (6.5.2).

Similarly, if $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ are absolutely convergent series of complex numbers, then

$$(6.5.12) \quad \sum_{j,l \geq 0} |a_j| |b_l| = \left(\sum_{j=0}^{\infty} |a_j| \right) \left(\sum_{l=0}^{\infty} |b_l| \right).$$

This means that $a_j b_l$ is a summable complex-valued function of (j, l) , so that (6.5.11) can be defined as in Section 11.8. In this situation, (6.5.11) can be expressed in terms of iterated sums, as in Section 11.16. This can be used to get that $\sum_{n=0}^{\infty} c_n$ converges absolutely, and that both sides of (6.5.2) are equal to (6.5.11).

6.6 The complex exponential function

If z is a complex number, then put

$$(6.6.1) \quad E(z) = \sum_{j=0}^{\infty} \frac{z^j}{j!},$$

as in (25) on p178 of [210]. It is easy to see that the series on the right converges absolutely, using the ratio test. One can also use the comparison test more directly, by comparing this series with a convergent geometric series. It follows that this defines a continuous complex-valued function on the complex plane, as in Subsection 6.1.3.

6.6.1 Exponentials of sums

If w is another complex number and n is a nonnegative integer, then

$$(6.6.2) \quad (z + w)^n = \sum_{j=0}^n \binom{n}{j} z^j w^{n-j},$$

by the binomial theorem, where

$$(6.6.3) \quad \binom{n}{j} = \frac{n!}{j!(n-j)!}$$

is the usual binomial coefficient. This implies that

$$(6.6.4) \quad E(z + w) = \sum_{n=0}^{\infty} \frac{(z + w)^n}{n!} = \sum_{n=0}^{\infty} \left(\sum_{j=0}^n \frac{z^j}{j!} \frac{w^{n-j}}{(n-j)!} \right).$$

The right side is the same as the Cauchy product of the series corresponding to $E(z)$ and $E(w)$, as in Section 6.5. It follows that

$$(6.6.5) \quad E(z + w) = E(z) E(w),$$

as in (26) on p178 of [210], because these series converge absolutely.

6.6.2 Some more properties of $E(z)$

In particular,

$$(6.6.6) \quad E(z) E(-z) = E(z - z) = E(0) = 1,$$

as in (27) on p178 of [210], so that $E(z) \neq 0$, and

$$(6.6.7) \quad E(z)^{-1} = E(-z).$$

Of course, $E(x) \in \mathbf{R}$ when $x \in \mathbf{R}$, and it is easy to see that $E(x)$ is positive and strictly increasing for $x \geq 0$. The same properties hold when $x \leq 0$, because of (6.6.7). In fact, $E(x) \rightarrow +\infty$ as $x \rightarrow +\infty$, and thus $E(x) \rightarrow 0$ as $x \rightarrow -\infty$, as on p179 of [210].

One may use $E(x)$ as the definition of e^x when $x \in \mathbf{R}$, as on p179 of [210], or show that it is equivalent to other standard definitions. More precisely, one may define e to be $E(1)$, as in Definition 3.30 on p63 of [210]. Similarly, we may use $E(z)$ as the definition of the *complex exponential function* $e^z = \exp(z)$ for $z \in \mathbf{C}$, as on p1 of [209]. We shall use this notation from now on.

6.6.3 Exponentials and complex conjugation

If $z = x + iy$, $x, y \in \mathbf{R}$, then the *complex conjugate* of z is the complex number defined by

$$(6.6.8) \quad \bar{z} = x - iy.$$

If w is another complex number, then

$$(6.6.9) \quad \overline{(w + z)} = \bar{w} + \bar{z}$$

and

$$(6.6.10) \quad \overline{(wz)} = \bar{w} \bar{z}.$$

It is easy to see that

$$(6.6.11) \quad z \bar{z} = x^2 + y^2 = |z|^2,$$

where $|z|$ is the usual absolute value of z . The well-known fact that

$$(6.6.12) \quad |wz| = |z| |w|$$

follows from (6.6.10) and (6.6.11). Clearly

$$(6.6.13) \quad \overline{(\bar{z})} = z$$

and

$$(6.6.14) \quad |\bar{z}| = |z|.$$

One can check that

$$(6.6.15) \quad \overline{(\exp z)} = \exp(\bar{z}).$$

This implies that

$$(6.6.16) \quad \begin{aligned} |\exp z|^2 &= (\exp z) \overline{(\exp z)} = (\exp z) (\exp \bar{z}) \\ &= \exp(z + \bar{z}) = \exp(2 \operatorname{Re} z), \end{aligned}$$

where $\operatorname{Re} z$ is the real part of z .

In particular, $|\exp(iy)| = 1$ for every $y \in \mathbf{R}$. In fact, *Euler's identity* states that

$$(6.6.17) \quad \exp(iy) = \cos y + i \sin y,$$

as in (46) on p182 of [210], and (5) on p2 of [209].

6.6.4 Differentiating $\exp(at)$

If $a \in \mathbf{C}$, then $\exp(at)$ may be considered as a complex-valued function of $t \in \mathbf{R}$. This function is differentiable on \mathbf{R} , with derivative equal to $a \exp(at)$. This can be seen by differentiating the corresponding power series termwise, as in Subsection 6.3.3. Alternatively, one can reduce to taking the derivative at 0, using (6.6.5), as on p179 of [210] and p2 of [209]. More precisely, $\exp z$ is a holomorphic function on the complex plane, whose derivative is equal to itself.

6.7 Abel summability

Let $\sum_{j=0}^{\infty} a_j$ be an infinite series of real or complex numbers. Suppose that for each real number r with $0 \leq r < 1$, we have that

$$(6.7.1) \quad \sum_{j=0}^{\infty} |a_j| r^j \text{ converges.}$$

This happens when $\{a_j\}_{j=0}^{\infty}$ is bounded, for instance, by the comparison test, because $\sum_{j=0}^{\infty} r^j$ is a convergent geometric series. Put

$$(6.7.2) \quad A(r) = \sum_{j=0}^{\infty} a_j r^j$$

for each such r .

$$(6.7.3) \quad \text{Put} \quad \lim_{r \rightarrow 1-} A(r) = A,$$

when the limit on the left side exists. In this case, we say that $\sum_{j=0}^{\infty} a_j$ is *Abel summable*, with Abel sum equal to A .

If $\sum_{j=0}^{\infty} a_j$ converges in the usual sense, then a famous theorem of Abel states that $\sum_{j=0}^{\infty} a_j$ is Abel summable, with

$$(6.7.4) \quad A = \sum_{j=0}^{\infty} a_j.$$

Of course, if $\sum_{j=0}^{\infty} a_j$ converges in the usual sense, then $\{a_j\}_{j=0}^{\infty}$ converges to 0, and is thus bounded, so that (6.7.1) converges when $0 \leq r < 1$.

If $\sum_{j=0}^{\infty} a_j$ converges absolutely, then one can get Abel summability with Abel sum as in (6.7.4) from the remarks in Subsection 6.1.2.

6.7.1 The proof of Abel's theorem

To see this, let

$$(6.7.5) \quad s_n = \sum_{j=0}^n a_j$$

be the n th partial sum of $\sum_{j=0}^{\infty} a_j$ for each nonnegative integer n , and put $s_{-1} = 0$, so that

$$(6.7.6) \quad a_n = s_n - s_{n-1}$$

for each $n \geq 0$. If $0 \leq r < 1$, then

$$(6.7.7) \quad A(r) = \sum_{j=0}^{\infty} s_j r^j - \sum_{j=0}^{\infty} s_{j-1} r^j,$$

where the series on the right converge absolutely because $\{s_n\}_{n=0}^{\infty}$ converges and is thus bounded.

Observe that

$$(6.7.8) \quad \sum_{j=0}^{\infty} s_{j-1} r^j = \sum_{j=0}^{\infty} s_j r^{j+1} = r \sum_{j=0}^{\infty} s_j r^j,$$

using $s_{-1} = 0$ in the first step. It follows that

$$(6.7.9) \quad A(r) = (1 - r) \sum_{j=0}^{\infty} s_j r^j.$$

6.7.2 Estimating $A(r) - s$

Put

$$(6.7.10) \quad s = \sum_{j=0}^{\infty} a_j,$$

so that $\{s_n\}_{n=0}^{\infty}$ converges to s , by hypothesis. Note that

$$(6.7.11) \quad (1 - r) \sum_{j=0}^{\infty} r^j = 1$$

when $0 \leq r < 1$. This means that

$$(6.7.12) \quad A(r) - s = (1 - r) \sum_{j=0}^{\infty} (s_j - s) r^j.$$

It follows that

$$(6.7.13) \quad |A(r) - s| \leq (1 - r) \sum_{j=0}^{\infty} |s_j - s| r^j$$

when $0 \leq r < 1$.

6.7.3 The limit as $r \rightarrow 1-$

If N is any nonnegative integer, then we get that $|A(r) - s|$ is less than or equal to the sum of

$$(6.7.14) \quad (1-r) \sum_{j=0}^N |s_j - s| r^j$$

and

$$(6.7.15) \quad (1-r) \sum_{j=N+1}^{\infty} |s_j - s| r^j.$$

One can check that (6.7.15) tends to 0 as $N \rightarrow \infty$, uniformly over $0 \leq r < 1$, because $\{s_j\}_{j=0}^{\infty}$ converges to s , by hypothesis, and using (6.7.11). If N is fixed, then (6.7.14) tends to 0 as $r \rightarrow 1-$. One can use this to get that $A(r)$ tends to s as $r \rightarrow 1-$.

6.8 Abel summability and Cauchy products

Let $\sum_{j=0}^{\infty} a_j$, $\sum_{l=0}^{\infty} b_l$ be infinite series of real or complex numbers, and let $\sum_{n=0}^{\infty} c_n$ be their Cauchy product, as in Section 6.5. Suppose that (6.7.1) holds when $0 \leq r < 1$, and similarly that

$$(6.8.1) \quad \sum_{l=0}^{\infty} |b_l| r^l \text{ converges}$$

when $0 \leq r < 1$. Note that $\sum_{n=0}^{\infty} c_n r^n$ is the Cauchy product of $\sum_{j=0}^{\infty} a_j r^j$ and $\sum_{l=0}^{\infty} b_l r^l$, as in Section 6.5. It follows that

$$(6.8.2) \quad \sum_{n=0}^{\infty} |c_n| r^n \text{ converges}$$

when $0 \leq r < 1$, as in Subsection 6.5.2.

Let $A(r)$ be as in (6.7.2) again, and similarly put

$$(6.8.3) \quad B(r) = \sum_{l=0}^{\infty} b_l r^l$$

and

$$(6.8.4) \quad C(r) = \sum_{n=0}^{\infty} c_n r^n$$

when $0 \leq r < 1$. Under these conditions, we have that

$$(6.8.5) \quad C(r) = A(r) B(r)$$

when $0 \leq r < 1$, as in Subsection 6.5.2.

6.8.1 Abel summability conditions

Suppose that $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ are Abel summable, let A be as in (6.7.3), and similarly put

$$(6.8.6) \quad \lim_{r \rightarrow 1-} B(r) = B.$$

It follows that $\sum_{n=0}^{\infty} c_n$ is Abel summable, with

$$(6.8.7) \quad \lim_{r \rightarrow 1-} C(r) = C = AB,$$

because of (6.8.5).

6.8.2 A corollary about Cauchy products

In particular, this works when $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$ converge in the usual sense, because of Abel's theorem, as in the previous section. If $\sum_{n=0}^{\infty} c_n$ converges in the usual sense as well, then (6.8.7) implies that $\sum_{n=0}^{\infty} c_n$ is equal to the product of $\sum_{j=0}^{\infty} a_j$ and $\sum_{l=0}^{\infty} b_l$.

This corresponds to Exercise 6 on p256 of [93], and Theorem 3.51 on p75 of [210]. More precisely, the proof of the latter is discussed on p175 of [210].

6.9 Cesaro means

Let $\{s_n\}_{n=0}^{\infty}$ be an infinite sequence of real or complex numbers, and put

$$(6.9.1) \quad \sigma_n = \frac{1}{n+1} \sum_{j=0}^n s_j$$

for each nonnegative integer n . If $\{\sigma_n\}_{n=0}^{\infty}$ converges to a real or complex number σ , then we say that $\{s_n\}_{n=0}^{\infty} (C, 1)$ converges to σ , as on p55 of [93].

If $\{s_n\}_{n=0}^{\infty}$ converges to a real or complex number s , as appropriate, then one can show that $\{s_n\}_{n=0}^{\infty} (C, 1)$ converges to s . This corresponds to Theorem 2.11 B on p56 of [93], and to part (a) of Exercise 14 on p80 of [210].

If $\{s_n\}_{n=0}^{\infty} (C, 1)$ converges, then one can show that

$$(6.9.2) \quad \lim_{n \rightarrow \infty} \frac{s_n}{n+1} = 0,$$

as in Exercise 4 on p64 of [93].

6.9.1 Cesaro summability of infinite series

Let $\sum_{j=0}^{\infty} a_j$ be an infinite series of real or complex numbers, and let $s_n = \sum_{j=0}^n a_j$ be the corresponding sequence of partial sums, for $n \geq 0$. If $\{s_n\}_{n=0}^{\infty} (C, 1)$ converges to a real or complex number σ , then we say that $\sum_{j=0}^{\infty} a_j$ is $(C, 1)$ summable with $(C, 1)$ sum σ , as in Definition 3.9 A on p91 of [93]. If

$\sum_{j=0}^{\infty} a_j$ converges in the usual sense, then $\sum_{j=0}^{\infty} a_j$ is $(C, 1)$ summable with the same sum, as before.

If $\sum_{j=0}^{\infty} a_j$ is $(C, 1)$ summable, then one show that

$$(6.9.3) \quad \{a_j\}_{j=0}^{\infty} (C, 1) \text{ converges to } 0,$$

as in Exercise 5 on p93 of [93]. In particular, this implies that

$$(6.9.4) \quad \lim_{j \rightarrow \infty} \frac{a_j}{j+1} = 0,$$

as before.

Note that (6.9.4) implies that (6.7.1) holds for each $0 \leq r < 1$. This means that $A(r)$ may be defined as in (6.7.2) when $0 \leq r < 1$.

If $\sum_{j=0}^{\infty} a_j$ is $(C, 1)$ summable, then it is well known that $\sum_{j=0}^{\infty} a_j$ is Abel summable, with the same sum. This corresponds to Theorem 9.6 E on p252 of [93].

6.10 Comparing norms on \mathbf{R}^n , \mathbf{C}^n

Let n be a positive integer, and let

$$(6.10.1) \quad e_j = (e_{j,1}, \dots, e_{j,n})$$

be the j th standard basis vector in \mathbf{R}^n for each $j = 1, \dots, n$. Thus

$$(6.10.2) \quad \begin{aligned} e_{j,l} &= 1 && \text{when } j = l \\ &= 0 && \text{when } j \neq l. \end{aligned}$$

If $v = (v_1, \dots, v_n) \in \mathbf{R}^n$, then v can be expressed as

$$(6.10.3) \quad v = \sum_{j=1}^n v_j e_j,$$

where the right side is a linear combination of elements of \mathbf{R}^n . The e_j 's may also be considered as standard basis vectors in \mathbf{C}^n . If $v = (v_1, \dots, v_n) \in \mathbf{C}^n$, then v can be expressed as in (6.10.3) again, where now the right side is a linear combination of elements of \mathbf{C}^n .

6.10.1 Using the standard basis

Let N be a norm on \mathbf{R}^n or \mathbf{C}^n , as in Subsection 1.3.1 and Section 1.4. If $v = (v_1, \dots, v_n)$ is an element of \mathbf{R}^n or \mathbf{C}^n , as appropriate, then

$$(6.10.4) \quad N(v) = N\left(\sum_{j=1}^n v_j e_j\right),$$

by (6.10.3). This implies that

$$(6.10.5) \quad N(v) \leq \sum_{j=1}^n N(v_j e_j) = \sum_{j=1}^n |v_j| N(e_j),$$

using the triangle inequality for N in the first step, and the homogeneity condition for N in the second step. Note that this argument also works for seminorms instead of norms, where seminorms are defined as in Subsection A.6.1.

6.10.2 Comparison with the norm $\|\cdot\|_1$

It follows that

$$(6.10.6) \quad N(v) \leq \left(\max_{1 \leq l \leq n} N(e_l) \right) \sum_{j=1}^n |v_j| = \left(\max_{1 \leq l \leq n} N(e_l) \right) \|v\|_1,$$

where $\|v\|_1$ is as in (1.3.6) or (1.4.2), as appropriate. Remember that

$$(6.10.7) \quad d_N(v, w) = N(v - w)$$

is the metric on \mathbf{R}^n or \mathbf{C}^n , as appropriate, associated to N , as in (1.3.9) and (1.4.4). Similarly, we let

$$(6.10.8) \quad d_1(v, w) = \|v - w\|_1$$

be the metric on \mathbf{R}^n or \mathbf{C}^n , as appropriate, associated to $\|\cdot\|_1$, as in (1.3.11) and (1.4.6). Thus

$$(6.10.9) \quad d_N(v, w) \leq \left(\max_{1 \leq l \leq n} N(e_l) \right) d_1(v, w)$$

for every v, w in \mathbf{R}^n or \mathbf{C}^n , as appropriate. In particular, the identity mapping on \mathbf{R}^n or \mathbf{C}^n , as appropriate, is Lipschitz with respect to (6.10.8) on the domain and (6.10.7) on the range.

6.10.3 Comparison with the norm $\|\cdot\|_2$

Observe that

$$(6.10.10) \quad N(v) \leq \left(\sum_{l=1}^n N(e_l)^2 \right)^{1/2} \left(\sum_{j=1}^n |v_j|^2 \right)^{1/2}$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, using the Cauchy–Schwarz inequality (1.3.13) on the right side of (6.10.5). Equivalently, this means that

$$(6.10.11) \quad N(v) \leq \left(\sum_{l=1}^n N(e_l)^2 \right)^{1/2} \|v\|_2,$$

where $\|v\|_2$ is the standard Euclidean norm of v , as in (1.3.5) or (1.4.1), as appropriate. Remember that

$$(6.10.12) \quad d_2(v, w) = \|v - w\|_2$$

is the standard Euclidean metric on \mathbf{R}^n or \mathbf{C}^n , as appropriate, as in (1.3.10) and (1.4.5). It follows that

$$(6.10.13) \quad d_N(v, w) \leq \left(\sum_{l=1}^n N(e_l)^2 \right)^{1/2} d_2(v, w)$$

for every v, w in \mathbf{R}^n or \mathbf{C}^n , as appropriate. This implies that the identity mapping on \mathbf{R}^n or \mathbf{C}^n , as appropriate, is Lipschitz with respect to (6.10.12) on the domain and (6.10.7) on the range.

6.10.4 Comparison with the norm $\|\cdot\|_\infty$

Using (6.10.5), we also get that

$$(6.10.14) \quad N(v) \leq \left(\sum_{j=1}^n N(e_j) \right) \max_{1 \leq l \leq n} |v_l| = \left(\sum_{j=1}^n N(e_j) \right) \|v\|_\infty$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. Here $\|v\|_\infty$ is as in (1.3.7) or (1.4.3), as appropriate. Let

$$(6.10.15) \quad d_\infty(v, w) = \|v - w\|_\infty$$

be the metric on \mathbf{R}^n or \mathbf{C}^n , as appropriate, associated to $\|\cdot\|_\infty$ as in (1.3.12) and (1.4.7). Note that

$$(6.10.16) \quad d_N(v, w) \leq \left(\sum_{j=1}^n N(e_j) \right) d_\infty(v, w)$$

for every v, w in \mathbf{R}^n or \mathbf{C}^n , as appropriate. Hence the identity mapping on \mathbf{R}^n or \mathbf{C}^n , as appropriate, is Lipschitz with respect to (6.10.15) on the domain and (6.10.7) on the range.

6.11 Another comparison

Let N be a norm on \mathbf{R}^n or \mathbf{C}^n for some $n \in \mathbf{Z}_+$ again. Observe that

$$(6.11.1) \quad N(v) \leq N(w) + N(v - w)$$

and

$$(6.11.2) \quad N(w) \leq N(v) + N(v - w)$$

for every v, w in \mathbf{R}^n or \mathbf{C}^n , as appropriate. This implies that

$$(6.11.3) \quad |N(v) - N(w)| = \max(N(v) - N(w), N(w) - N(v)) \leq N(v - w),$$

for every v, w in \mathbf{R}^n or \mathbf{C}^n , as appropriate. More precisely, the left side of (6.11.3) is the usual absolute value of $N(v) - N(w)$, as a real number.

6.11.1 Continuity of N

Remember that

$$(6.11.4) \quad N(v) \leq C \|v\|_2$$

for some nonnegative real number C and every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, as in (6.10.11). Combining this with (6.11.3), we get that

$$(6.11.5) \quad |N(v) - N(w)| \leq C \|v - w\|_2 = C d_2(v, w)$$

for every v, w in \mathbf{R}^n or \mathbf{C}^n , as appropriate.

Thus N is Lipschitz as a real-valued function on \mathbf{R}^n or \mathbf{C}^n , as appropriate, using the corresponding Euclidean metrics on the domain and range. In particular, N is continuous. This argument works for seminorms instead of norms as well.

6.11.2 Using the extreme value theorem

The extreme value theorem implies that there is a $u \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, such that $\|u\|_2 = 1$ and

$$(6.11.6) \quad N(u) \leq N(w)$$

for every $w \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, with $\|w\|_2 = 1$. This uses the fact that the unit sphere in \mathbf{R}^n or \mathbf{C}^n , as appropriate, is compact with respect to the Euclidean metric.

Put

$$(6.11.7) \quad c = N(u),$$

and note that $c > 0$, because N is a norm. Let us check that

$$(6.11.8) \quad c \|v\|_2 \leq N(v)$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n . Of course, (6.11.8) is trivial when $v = 0$, and so we may suppose that $v \neq 0$.

In this case, $\|v\|_2 > 0$, and

$$(6.11.9) \quad w = \frac{v}{\|v\|_2}$$

satisfies $\|w\|_2 = 1$. Thus

$$(6.11.10) \quad N(v)/\|v\|_2 = N(v/\|v\|_2) = N(w) \geq N(u) = c,$$

using (6.11.6) in the third step. This implies (6.11.8), as desired.

6.11.3 Another comparison with $\|\cdot\|_2$

Put $C' = 1/c$, so that (6.11.8) is the same as saying that

$$(6.11.11) \quad \|v\|_2 \leq C' N(v)$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. It follows that

$$(6.11.12) \quad d_2(v, w) = \|v - w\|_2 \leq C' N(v - w) = C' d_N(v, w)$$

for every v, w in \mathbf{R}^n or \mathbf{C}^n , as appropriate.

This means that the identity mapping on \mathbf{R}^n or \mathbf{C}^n , as appropriate, is Lipschitz with respect to the metric $d_N(v, w)$ associated to N on the domain, and the standard Euclidean metric on the range. Of course, there are analogous arguments using $\|\cdot\|_1$ or $\|\cdot\|_\infty$ here, instead of the Euclidean norm $\|\cdot\|_2$.

6.12 The standard inner products on \mathbf{R}^n , \mathbf{C}^n

Let n be a positive integer, and let $x, y \in \mathbf{R}^n$ be given. Thus

$$(6.12.1) \quad \langle x, y \rangle = \langle x, y \rangle_{\mathbf{R}^n} = \sum_{j=1}^n x_j y_j$$

is defined as a real number. This is the standard *inner product* on \mathbf{R}^n . Note that

$$(6.12.2) \quad \langle x, x \rangle_{\mathbf{R}^n} = \sum_{j=1}^n x_j^2 = \|x\|_2^2,$$

where $\|x\|_2$ is the standard Euclidean norm of x , as in (1.3.5). We also have that

$$(6.12.3) \quad |\langle x, y \rangle_{\mathbf{R}^n}| = \left| \sum_{j=1}^n x_j y_j \right| \leq \sum_{j=1}^n |x_j| |y_j| \leq \|x\|_2 \|y\|_2,$$

using the Cauchy-Schwarz inequality (1.3.13) in the third step.

6.12.1 Using complex numbers

Let $z = x + iy$ and $w = u + iv$ be complex numbers, where x, y, u , and v are real numbers. The real part of $w\bar{z}$ is given by

$$(6.12.4) \quad \operatorname{Re}(w\bar{z}) = ux + vy.$$

As usual, w and z correspond to (u, v) and (x, y) , respectively, as elements of \mathbf{R}^2 . The right side of (6.12.4) is the same as the inner product of (u, v) and (x, y) as elements of \mathbf{R}^2 .

6.12.2 The standard (Hermitian) inner product on \mathbf{C}^n

If $w, z \in \mathbf{C}^n$, then

$$(6.12.5) \quad \langle w, z \rangle = \langle w, z \rangle_{\mathbf{C}^n} = \sum_{j=1}^n w_j \bar{z}_j$$

is defined as a complex number. This is the standard inner product on \mathbf{C}^n . Observe that

$$(6.12.6) \quad \overline{\langle w, z \rangle_{\mathbf{C}^n}} = \overline{\left(\sum_{j=1}^n w_j \bar{z}_j \right)} = \sum_{j=1}^n \bar{w}_j z_j = \langle z, w \rangle_{\mathbf{C}^n}.$$

As before,

$$(6.12.7) \quad \langle z, z \rangle_{\mathbf{C}^n} = \sum_{j=1}^n z_j \bar{z}_j = \sum_{j=1}^n |z_j|^2 = \|z\|_2^2,$$

where $\|z\|_2$ is the standard Euclidean norm of z , as in (1.4.1). Moreover,

$$(6.12.8) \quad |\langle w, z \rangle_{\mathbf{C}^n}| = \left| \sum_{j=1}^n w_j \bar{z}_j \right| \leq \sum_{j=1}^n |w_j| |z_j| \leq \|w\|_2 \|z\|_2,$$

using the Cauchy–Schwarz inequality (1.3.13) in the third step.

6.13 Sums and inner products

Let X be a nonempty set, and let f, g be real-valued functions on X with finite support in X . Put

$$(6.13.1) \quad \langle f, g \rangle = \langle f, g \rangle_{c_{00}(X, \mathbf{R})} = \sum_{x \in X} f(x) g(x),$$

where the sum on the right reduces to a finite sum of real numbers, as in Subsection 1.6.1. This may be considered as the standard inner product on $c_{00}(X, \mathbf{R})$. Of course,

$$(6.13.2) \quad \langle f, f \rangle_{c_{00}(X, \mathbf{R})} = \sum_{x \in X} f(x)^2 = \|f\|_2^2,$$

where $\|f\|_2$ is as in Subsection 1.6.1. Observe that

$$(6.13.3) \quad |\langle f, g \rangle_{c_{00}(X, \mathbf{R})}| = \left| \sum_{x \in X} f(x) g(x) \right| \leq \sum_{x \in X} |f(x)| |g(x)| \leq \|f\|_2 \|g\|_2,$$

by the Cauchy–Schwarz inequality.

6.13.1 Another Hermitian inner product

Similarly, if f, g are complex-valued functions on X with finite support, then

$$(6.13.4) \quad \langle f, g \rangle = \langle f, g \rangle_{c_{00}(X, \mathbf{C})} = \sum_{x \in X} f(x) \overline{g(x)}$$

reduces to a finite sum of complex numbers. This may be considered as the standard inner product on $c_{00}(X, \mathbf{C})$. As before,

$$(6.13.5) \quad \langle f, f \rangle_{c_{00}(X, \mathbf{C})} = \sum_{x \in X} |f(x)|^2 = \|f\|_2^2$$

for every $f \in c_{00}(X, \mathbf{C})$. Using the Cauchy–Schwarz inequality again, we get that

$$(6.13.6) \quad |\langle f, g \rangle_{c_{00}(X, \mathbf{C})}| = \left| \sum_{x \in X} f(x) \overline{g(x)} \right| \leq \sum_{x \in X} |f(x)| |g(x)| \leq \|f\|_2 \|g\|_2$$

for every $f, g \in c_{00}(X, \mathbf{C})$. We also have that

$$(6.13.7) \quad \overline{\langle f, g \rangle_{c_{00}(X, \mathbf{C})}} = \overline{\left(\sum_{x \in X} f(x) \overline{g(x)} \right)} = \sum_{x \in X} g(x) \overline{f(x)} = \langle g, f \rangle_{c_{00}(X, \mathbf{C})}$$

for every $f, g \in c_{00}(X, \mathbf{C})$.

6.13.2 Inner products on $\ell^2(\mathbf{Z}_+, \mathbf{R})$, $\ell^2(\mathbf{Z}_+, \mathbf{C})$

Suppose now that $f, g \in \ell^2(\mathbf{Z}_+, \mathbf{R})$, and remember that $f, g \in \ell^1(\mathbf{Z}_+, \mathbf{R})$, as in Section 2.3. In this case, we put

$$(6.13.8) \quad \langle f, g \rangle = \langle f, g \rangle_{\ell^2(\mathbf{Z}_+, \mathbf{R})} = \sum_{j=1}^{\infty} f(j) g(j),$$

where the right side converge absolutely. This may be considered as the standard inner product on $\ell^2(\mathbf{Z}_+, \mathbf{R})$. Note that

$$(6.13.9) \quad \langle f, f \rangle_{\ell^2(\mathbf{Z}_+, \mathbf{R})} = \sum_{j=1}^{\infty} f(j)^2 = \|f\|_2^2,$$

where $\|f\|_2$ is as in Section 2.3. The version of the Cauchy–Schwarz inequality mentioned in Section 2.3 implies that

$$(6.13.10) \quad |\langle f, g \rangle_{\ell^2(\mathbf{Z}_+, \mathbf{R})}| = \left| \sum_{j=1}^{\infty} f(j) g(j) \right| \leq \sum_{j=1}^{\infty} |f(j)| |g(j)| \leq \|f\|_2 \|g\|_2.$$

Let $f, g \in \ell^2(\mathbf{Z}_+, \mathbf{C})$ be given, so that $|f|, |g| \in \ell^1(\mathbf{Z}_+, \mathbf{C})$, and hence $f \overline{g}$ is an element of $\ell^1(\mathbf{Z}_+, \mathbf{C})$. Put

$$(6.13.11) \quad \langle f, g \rangle = \langle f, g \rangle_{\ell^2(\mathbf{Z}_+, \mathbf{C})} = \sum_{j=1}^{\infty} f(j) \overline{g(j)},$$

which may be considered as the standard inner product on $\ell^2(\mathbf{Z}_+, \mathbf{C})$. As usual,

$$(6.13.12) \quad \langle f, f \rangle_{\ell^2(\mathbf{Z}_+, \mathbf{C})} = \sum_{j=1}^{\infty} |f(j)|^2 = \|f\|_2^2.$$

We also have that

$$(6.13.13) \quad |\langle f, g \rangle_{\ell^2(\mathbf{Z}_+, \mathbf{C})}| = \left| \sum_{j=1}^{\infty} f(j) \overline{g(j)} \right| \leq \sum_{j=1}^{\infty} |f(j)| |g(j)| \leq \|f\|_2 \|g\|_2,$$

by the version of the Cauchy–Schwarz inequality in Section 2.3. Moreover,

$$(6.13.14) \quad \overline{\langle f, g \rangle_{\ell^2(\mathbf{Z}_+, \mathbf{C})}} = \overline{\left(\sum_{j=1}^{\infty} f(j) \overline{g(j)} \right)} = \sum_{j=1}^{\infty} g(j) \overline{f(j)} = \langle g, f \rangle_{\ell^2(\mathbf{Z}_+, \mathbf{C})}.$$

6.14 Some integral inner products

Let a, b be real numbers with $a < b$, and let α be a monotonically increasing real-valued function on $[a, b]$. If f, g are continuous real-valued functions on $[a, b]$, then

$$(6.14.1) \quad \langle f, g \rangle = \langle f, g \rangle_{\alpha} = \int_a^b f(x) g(x) d\alpha(x)$$

is defined as a real number, using a Riemann–Stieltjes integral on the right side.

Clearly

$$(6.14.2) \quad \langle f, f \rangle_{\alpha} = \int_a^b f(x)^2 d\alpha(x) = \|f\|_{2, \alpha}^2,$$

where $\|f\|_{2, \alpha}$ is as in (3.5.1). Observe that

$$(6.14.3) \quad \begin{aligned} |\langle f, g \rangle_{\alpha}| &= \left| \int_a^b f(x) g(x) d\alpha(x) \right| \\ &\leq \int_a^b |f(x)| |g(x)| d\alpha(x) \leq \|f\|_{2, \alpha} \|g\|_{2, \alpha}, \end{aligned}$$

by the integral version (3.5.8) of the Cauchy–Schwarz inequality for Riemann–Stieltjes integrals. Of course, (6.14.1) is symmetric in f and g , and linear in each of f and g , because of linearity of the integral.

If α is strictly increasing on $[a, b]$, then (6.14.2) is positive when $f \not\equiv 0$ on $[a, b]$, as before. In this case, (6.14.1) defines an inner product on the space $C([a, b], \mathbf{R})$ of continuous real-valued functions on $[a, b]$.

6.14.1 Some inner products on $C([a, b], \mathbf{C})$

If f and g are continuous complex-valued functions on $[a, b]$, then

$$(6.14.4) \quad \langle f, g \rangle = \langle f, g \rangle_{\alpha} = \int_a^b f(x) \overline{g(x)} d\alpha(x)$$

is defined as a complex number, using a Riemann–Stieltjes integral on the right side again. In this situation, we also have that

$$(6.14.5) \quad \langle f, f \rangle_{\alpha} = \int_a^b |f(x)|^2 d\alpha(x) = \|f\|_{2, \alpha}^2,$$

where $\|f\|_{2,\alpha}$ is as in (3.5.1). As before,

$$(6.14.6) \quad |\langle f, g \rangle_\alpha| = \left| \int_a^b f(x) \overline{g(x)} d\alpha(x) \right| \leq \int_a^b |f(x)| |g(x)| d\alpha(x) \leq \|f\|_{2,\alpha} \|g\|_{2,\alpha},$$

using (3.5.8) in the third step. Note that

$$(6.14.7) \quad \overline{\langle f, g \rangle_\alpha} = \overline{\left(\int_a^b f(x) \overline{g(x)} d\alpha(x) \right)} = \int_a^b g(x) \overline{f(x)} d\alpha(x) = \langle g, f \rangle_\alpha.$$

It is easy to see that (6.14.4) is linear in f , because of linearity of the integral. Similarly, (6.14.4) is conjugate-linear in g , which is to say that it is additive in g , while multiplying g by a complex number t corresponds to multiplying (6.14.4) by \bar{t} .

If α is strictly increasing on $[a, b]$, then (6.14.5) is positive when $f \neq 0$ on $[a, b]$, and (6.14.4) defines an inner product on the space $C([a, b], \mathbf{C})$ of continuous complex-valued functions on $[a, b]$, as a vector space over the complex numbers. If $\alpha(x) = x$ for every $x \in [a, b]$, then (6.14.1) and (6.14.4) are called the standard integral inner products on $C([a, b], \mathbf{R})$ and $C([a, b], \mathbf{C})$, respectively.

6.14.2 Some inner products on $C_{com}(\mathbf{R}, \mathbf{R})$

Now let α be a monotonically increasing real-valued function on \mathbf{R} , and let f, g be continuous real-valued functions on the real line with compact support. In this case,

$$(6.14.8) \quad \langle f, g \rangle = \langle f, g \rangle_\alpha = \int_{-\infty}^{\infty} f(x) g(x) d\alpha(x)$$

is defined as a real number, where the integral on the right is as in Section 3.8. As before,

$$(6.14.9) \quad \langle f, f \rangle_\alpha = \int_{-\infty}^{\infty} f(x)^2 d\alpha(x) = \|f\|_{2,\alpha}^2,$$

where the right side is as in Subsection 3.8.1. We also have that

$$(6.14.10) \quad |\langle f, g \rangle_\alpha| = \left| \int_{-\infty}^{\infty} f(x) g(x) d\alpha(x) \right| \leq \int_{-\infty}^{\infty} |f(x)| |g(x)| d\alpha(x) \leq \|f\|_{2,\alpha} \|g\|_{2,\alpha},$$

using the version of the Cauchy–Schwarz inequality mentioned in Subsection 3.8.3 in the third step. Clearly (6.14.8) is symmetric in f and g , and linear in each of f and g , because of the linearity of the integral.

If α is strictly increasing on \mathbf{R} , then (6.14.9) is positive when $f \neq 0$ on \mathbf{R} , as usual. This means that (6.14.8) defines an inner product on $C_{com}(\mathbf{R}, \mathbf{R})$.

6.14.3 Some inner products on $C_{com}(\mathbf{R}, \mathbf{C})$

Similarly, if f and g are continuous complex-valued functions on \mathbf{R} with compact support, then

$$(6.14.11) \quad \langle f, g \rangle = \langle f, g \rangle_\alpha = \int_{-\infty}^{\infty} f(x) \overline{g(x)} d\alpha(x)$$

is defined as a complex number, and

$$(6.14.12) \quad \langle f, f \rangle_\alpha = \int_{-\infty}^{\infty} |f(x)|^2 d\alpha(x) = \|f\|_{2,\alpha}^2.$$

In addition,

$$(6.14.13) \quad \begin{aligned} |\langle f, g \rangle_\alpha| &= \left| \int_{-\infty}^{\infty} f(x) \overline{g(x)} d\alpha(x) \right| \\ &\leq \int_{-\infty}^{\infty} |f(x)| |g(x)| d\alpha(x) \leq \|f\|_{2,\alpha} \|g\|_{2,\alpha} \end{aligned}$$

and

$$(6.14.14) \quad \begin{aligned} \overline{\langle f, g \rangle_\alpha} &= \overline{\left(\int_{-\infty}^{\infty} f(x) \overline{g(x)} d\alpha(x) \right)} \\ &= \int_{-\infty}^{\infty} g(x) \overline{f(x)} d\alpha(x) = \langle g, f \rangle_\alpha. \end{aligned}$$

Note that (6.14.11) is linear in f , because of the linearity of the integral, and conjugate-linear in g .

If α is strictly increasing on \mathbf{R} , then (6.14.12) is positive when $f \neq 0$ on \mathbf{R} , so that (6.14.11) defines an inner product on $C_{com}(\mathbf{R}, \mathbf{C})$, as a vector space over the complex numbers. If $\alpha(x) = x$ for all $x \in \mathbf{R}$, then (6.14.8) and (6.14.11) are called the standard integral inner products on $C_{com}(\mathbf{R}, \mathbf{R})$ and $C_{com}(\mathbf{R}, \mathbf{C})$, respectively.

Chapter 7

Matrix norms and Lipschitz conditions

7.1 Real and complex matrices

Let m and n be positive integers, and let us consider $m \times n$ matrices with entries in the real or complex numbers. Such a matrix may be denoted as

$$(7.1.1) \quad [a_{j,l}] = [a_{j,l}]_{j,l=1}^{m,n},$$

where $a_{j,l}$ is a real or complex number, as appropriate, for each $j = 1, \dots, m$ and $l = 1, \dots, n$. Let $M_{m,n}(\mathbf{R})$ and $M_{m,n}(\mathbf{C})$ be the spaces of $m \times n$ matrices with entries in \mathbf{R} and \mathbf{C} , respectively.

If $[a_{j,l}]$ and $[b_{j,l}]$ are $m \times n$ matrices with real or complex entries, then their sum is defined as an $m \times n$ matrix by adding the corresponding entries, so that

$$(7.1.2) \quad [a_{j,l}] + [b_{j,l}] = [a_{j,l} + b_{j,l}].$$

Similarly, if t is a real or complex number, as appropriate, then t times $[a_{j,l}]$ is defined as an $m \times n$ matrix by multiplying the entries of $[a_{j,l}]$ by t ,

$$(7.1.3) \quad t[a_{j,l}] = [ta_{j,l}].$$

This makes $M_{m,n}(\mathbf{R})$ and $M_{m,n}(\mathbf{C})$ into vector spaces over the real and complex numbers, respectively. Of course, one can also identify $M_{m,n}(\mathbf{R})$ and $M_{m,n}(\mathbf{C})$ with \mathbf{R}^{mn} and \mathbf{C}^{mn} , respectively, by listing the entries of an $m \times n$ matrix in a sequence with mn terms.

7.1.1 Norms on matrices

As usual, a nonnegative real-valued function N defined on $M_{m,n}(\mathbf{R})$ or $M_{m,n}(\mathbf{C})$ is said to be a *norm* if it satisfies the following three conditions. First,

$$(7.1.4) \quad N([a_{j,l}]) = 0$$

if and only if $[a_{j,l}] = 0$ as a matrix, which means that $a_{j,l} = 0$ for every $j = 1, \dots, m$ and $l = 1, \dots, n$. Second,

$$(7.1.5) \quad N(t[a_{j,l}]) = |t| N([a_{j,l}])$$

for every $[a_{j,l}] \in M_{m,n}(\mathbf{R})$ or $M_{m,n}(\mathbf{C})$ and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate. Third,

$$(7.1.6) \quad N([a_{j,l}] + [b_{j,l}]) \leq N([a_{j,l}]) + N([b_{j,l}])$$

for every $[a_{j,l}], [b_{j,l}] \in M_{m,n}(\mathbf{R})$ or $M_{m,n}(\mathbf{C})$, as appropriate. In this case,

$$(7.1.7) \quad d_N([a_{j,l}], [b_{j,l}]) = N([a_{j,l}] - [b_{j,l}]) = N([a_{j,l} - b_{j,l}])$$

defines a metric on $M_{m,n}(\mathbf{R})$ or $M_{m,n}(\mathbf{C})$, as appropriate.

If $M_{m,n}(\mathbf{R})$ and $M_{m,n}(\mathbf{C})$ are identified with $\mathbf{R}^{m \cdot n}$ and $\mathbf{C}^{m \cdot n}$, respectively, as before, then the definition of a norm on $M_{m,n}(\mathbf{R})$ or $M_{m,n}(\mathbf{C})$ corresponds exactly to the earlier definitions for $\mathbf{R}^{m \cdot n}$ and $\mathbf{C}^{m \cdot n}$ in Subsection 1.3.1 and Section 1.4, respectively. Similarly, the metric associated to a norm on $M_{m,n}(\mathbf{R})$ or $M_{m,n}(\mathbf{C})$ corresponds to the analogous notions for $\mathbf{R}^{m \cdot n}$ and $\mathbf{C}^{m \cdot n}$.

7.1.2 The Hilbert–Schmidt norm

Put

$$(7.1.8) \quad \|[a_{j,l}]\|_{HS} = \left(\sum_{j=1}^m \sum_{l=1}^n |a_{j,l}|^2 \right)^{1/2}$$

for every $[a_{j,l}] \in M_{m,n}(\mathbf{R})$ or $M_{m,n}(\mathbf{C})$. This is known as the *Hilbert–Schmidt norm* on $M_{m,n}(\mathbf{R})$ and $M_{m,n}(\mathbf{C})$. More precisely, (7.1.8) corresponds to the standard Euclidean norm on $\mathbf{R}^{m \cdot n}$ and $\mathbf{C}^{m \cdot n}$, using the identifications mentioned earlier. In particular, the fact that (7.1.8) defines a norm on $M_{m,n}(\mathbf{R})$ and $M_{m,n}(\mathbf{C})$ follows from the analogous statements for the standard Euclidean norms on $\mathbf{R}^{m \cdot n}$ and $\mathbf{C}^{m \cdot n}$. More precisely, the triangle inequality for (7.1.8) follows from the triangle inequality for the Euclidean norm, and the first two requirements of a norm can be verified directly.

7.2 Matrices and linear mappings

Let m and n be positive integers again. As usual, a mapping A from \mathbf{R}^n into \mathbf{R}^m is said to be *linear* if

$$(7.2.1) \quad A(v + w) = A(v) + A(w)$$

for every $v, w \in \mathbf{R}^n$, and

$$(7.2.2) \quad A(tv) = tA(v)$$

for every $v \in \mathbf{R}^n$ and $t \in \mathbf{R}$. Similarly, a mapping A from \mathbf{C}^n into \mathbf{C}^m is said to be (*complex*) *linear* if (7.2.1) holds for every $v, w \in \mathbf{C}^n$, and (7.2.2) holds for every $v \in \mathbf{C}^n$ and $t \in \mathbf{C}$.

7.2.1 Linear mappings associated to matrices

Let $[a_{j,l}]$ be an $m \times n$ matrix whose entries are real or complex numbers. If $v = (v_1, \dots, v_n) \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, then put

$$(7.2.3) \quad (A(v))_j = \sum_{l=1}^n a_{j,l} v_l$$

for each $j = 1, \dots, m$. This defines $A(v)$ as an element of \mathbf{R}^m or \mathbf{C}^m , as appropriate, where the j th coordinate of $A(v)$ is given by (7.2.3) for each $j = 1, \dots, m$. It is easy to see that A is linear as a mapping from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m , as appropriate. Conversely, it is well known and not too difficult to show that every linear mapping A from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m corresponds to a unique $m \times n$ matrix with real or complex entries, as appropriate, in this way.

7.2.2 Matrices associated to linear mappings

Let $e_k = (e_{k,1}, \dots, e_{k,n})$ be the k th standard basis vector in \mathbf{R}^n or \mathbf{C}^n for each $k = 1, \dots, n$. Thus $e_{k,l} = 1$ when $k = l$, and $e_{k,l} = 0$ when $k \neq l$. If A corresponds to $[a_{j,l}] \in M_{m,n}(\mathbf{R})$ or $M_{m,n}(\mathbf{C})$ as in (7.2.3), then

$$(7.2.4) \quad (A(e_k))_j = a_{j,k}$$

for every $j = 1, \dots, m$ and $k = 1, \dots, n$. If A is any linear mapping from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m , then one can use (7.2.4) to define an $m \times n$ matrix $[a_{j,l}]$ of real or complex numbers, as appropriate. Using this matrix, it is easy to see that (7.2.3) holds for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, because of linearity, and by expressing v as a linear combination of e_1, \dots, e_n .

7.2.3 Another description of the Hilbert–Schmidt norm

Let $[a_{j,l}] \in M_{m,n}(\mathbf{R})$ or $M_{m,n}(\mathbf{C})$ be given again, and let A be the corresponding linear mapping from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m , as in (7.2.3). Using (7.2.4), we get that

$$(7.2.5) \quad \|A(e_k)\|_2 = \left(\sum_{j=1}^m |a_{j,k}|^2 \right)^{1/2}$$

for every $k = 1, \dots, n$. More precisely, the left side of (7.2.5) refers to the standard Euclidean norm of $A(e_k)$ in \mathbf{R}^m or \mathbf{C}^m , as appropriate. It follows that the Hilbert–Schmidt norm of $[a_{j,l}]$ defined in the previous section can be given by

$$(7.2.6) \quad \|[a_{j,l}]\|_{HS} = \left(\sum_{k=1}^n \|A(e_k)\|_2^2 \right)^{1/2}.$$

7.2.4 A simple estimate using the Hilbert–Schmidt norm

Let $v = (v_1, \dots, v_n) \in \mathbf{R}^n$ or \mathbf{C}^n be given, as appropriate. Remember that

$$(7.2.7) \quad v = \sum_{k=1}^n v_k e_k,$$

where the right side is a linear combination of e_1, \dots, e_n in \mathbf{R}^n or \mathbf{C}^n , as appropriate. Using the linearity of A , we get that

$$(7.2.8) \quad A(v) = \sum_{k=1}^n v_k A(e_k),$$

where the right side is a linear combination of $A(e_1), \dots, A(e_k)$ in \mathbf{R}^m or \mathbf{C}^m , as appropriate. It follows that

$$(7.2.9) \quad \|A(v)\|_2 \leq \sum_{k=1}^n |v_k| \|A(e_k)\|_2,$$

where $\|\cdot\|_2$ refers to the standard Euclidean norm on \mathbf{R}^m or \mathbf{C}^m again, as appropriate. This implies that

$$(7.2.10) \quad \|A(v)\|_2 \leq \left(\sum_{k=1}^n |v_k|^2 \right)^{1/2} \| [a_{j,l}] \|_{HS},$$

using the Cauchy–Schwarz inequality on the right side of (7.2.9).

7.3 Some related estimates

Let m and n be positive integers, and let A be a linear mapping from \mathbf{R}^n into \mathbf{R}^m , or from \mathbf{C}^n into \mathbf{C}^m . Also let N be a norm on \mathbf{R}^m or \mathbf{C}^m , as appropriate.

If $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, then v can be expressed as a linear combination of the standard basis vectors e_1, \dots, e_n in \mathbf{R}^n or \mathbf{C}^n as in (7.2.7). Hence $A(v)$ can be expressed as in (7.2.8), so that

$$(7.3.1) \quad N(A(v)) \leq \sum_{k=1}^n |v_k| N(A(e_k)).$$

This corresponds to (7.2.9) when N is the standard Euclidean norm on \mathbf{R}^m or \mathbf{C}^m , as appropriate.

7.3.1 Estimates in terms of $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$

Let us now look at this in terms of various norms on \mathbf{R}^n or \mathbf{C}^n , as appropriate. Let us start with the standard Euclidean norm

$$(7.3.2) \quad \|v\|_2 = \left(\sum_{k=1}^n |v_k|^2 \right)^{1/2}$$

of v . As before, we can apply the Cauchy–Schwarz inequality to the right side of (7.3.1) to get that

$$(7.3.3) \quad N(A(v)) \leq \left(\sum_{k=1}^n N(A(e_k))^2 \right)^{1/2} \|v\|_2.$$

This is the same as (7.2.10) when N is the standard Euclidean norm on \mathbf{R}^m or \mathbf{C}^m , as appropriate.

Consider the norm

$$(7.3.4) \quad \|v\|_1 = \sum_{k=1}^n |v_k|$$

of v discussed in Subsection 1.3.2 and Section 1.4. It is easy to see that

$$(7.3.5) \quad N(A(v)) \leq \left(\max_{1 \leq k \leq n} N(A(e_k)) \right) \|v\|_1,$$

using (7.3.1).

Similarly, we can consider the norm

$$(7.3.6) \quad \|v\|_\infty = \max_{1 \leq k \leq n} |v_k|$$

of v discussed in Subsection 1.3.2 and Section 1.4 as well. Observe that

$$(7.3.7) \quad N(A(v)) \leq \left(\sum_{k=1}^n N(A(e_k)) \right) \|v\|_\infty,$$

by (7.3.1) again.

7.3.2 Using an arbitrary norm on \mathbf{R}^n or \mathbf{C}^n

Let N_0 be any norm on \mathbf{R}^n or \mathbf{C}^n , as appropriate. One can check that there is a nonnegative real number C such that

$$(7.3.8) \quad N(A(v)) \leq C N_0(v)$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. This follows from the previous remarks when $N_0(v)$ is given by (7.3.2), (7.3.4), or (7.3.6). Otherwise, one can use the fact that (7.3.2) is bounded by a constant multiple of $N_0(v)$, as in Subsection 6.11.3.

If $v, w \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, then

$$(7.3.9) \quad N(A(v) - A(w)) = N(A(v - w)) \leq C N_0(v - w).$$

This uses the linearity of A in the first step, and (7.3.8) in the second step. Thus

$$(7.3.10) \quad d_N(A(v), A(w)) \leq C d_{N_0}(v, w),$$

where d_{N_0} is the metric associated to N_0 on \mathbf{R}^n or \mathbf{C}^n , and d_N is the metric associated to N on \mathbf{R}^m or \mathbf{C}^m , as appropriate. Note that (7.3.10) implies (7.3.8), by taking $v = 0$.

7.4 Spaces of linear mappings

Let m and n be positive integers again. The space of linear mappings from \mathbf{R}^n into \mathbf{R}^m may be denoted $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$. Similarly, the space of (complex) linear mappings from \mathbf{C}^n into \mathbf{C}^m may be denoted $\mathcal{L}(\mathbf{C}^n, \mathbf{C}^m)$.

Let A, B be linear mappings from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m , and let t be a real or complex number, as appropriate. Thus tA can be defined as a mapping from \mathbf{R}^n into \mathbf{R}^m , or from \mathbf{C}^n into \mathbf{C}^m , as appropriate, by putting

$$(7.4.1) \quad (tA)(v) = tA(v)$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. Similarly, $A + B$ can be defined as a mapping from \mathbf{R}^n into \mathbf{R}^m , or from \mathbf{C}^n into \mathbf{C}^m , as appropriate, by putting

$$(7.4.2) \quad (A + B)(v) = A(v) + B(v)$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. It is easy to see that these are also linear as mappings from \mathbf{R}^n into \mathbf{R}^m , or from \mathbf{C}^n into \mathbf{C}^m , as appropriate.

As in Section 7.2, there are standard one-to-one correspondences between $M_{m,n}(\mathbf{R})$ and $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$, and between $M_{m,n}(\mathbf{C})$ and $\mathcal{L}(\mathbf{C}^n, \mathbf{C}^m)$. These correspondences are linear, in the sense that sums of matrices are associated to sums of linear mappings, and similarly for multiplication of matrices and linear mappings by real or complex numbers, as appropriate.

7.4.1 Norms on these spaces of linear mappings

A nonnegative real-valued function N on $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ or $\mathcal{L}(\mathbf{C}^n, \mathbf{C}^m)$ is said to be a *norm* if it satisfies the usual three conditions, as follows. First, if A is a linear mapping from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m , as appropriate, then

$$(7.4.3) \quad N(A) = 0$$

if and only if $A = 0$, which means that $A(v) = 0$ for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. Second, if A is a linear mapping from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then

$$(7.4.4) \quad N(tA) = |t|N(A).$$

Third, if A and B are linear mappings from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m , as appropriate, then

$$(7.4.5) \quad N(A + B) \leq N(A) + N(B).$$

Under these conditions,

$$(7.4.6) \quad d_N(A, B) = N(A - B)$$

defines a metric on $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ or $\mathcal{L}(\mathbf{C}^n, \mathbf{C}^m)$, as appropriate. Observe that norms on $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ and $\mathcal{L}(\mathbf{C}^n, \mathbf{C}^m)$ correspond exactly to norms on $M_{m,n}(\mathbf{R})$

and $M_{m,n}(\mathbf{C})$, respectively, using the correspondence between matrices and linear mappings described in Section 7.2. The metrics associated to these norms correspond to each other in the same way.

Let A be a linear mapping from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m again. Also let e_1, \dots, e_n be the standard basis vectors in \mathbf{R}^n or \mathbf{C}^n , as in Subsection 7.2.2, and let $\|\cdot\|_2$ be the standard Euclidean norm on \mathbf{R}^m or \mathbf{C}^m , as appropriate. Put

$$(7.4.7) \quad \|A\|_{HS} = \left(\sum_{k=1}^n \|A(e_k)\|_2^2 \right)^{1/2},$$

which corresponds exactly to the Hilbert–Schmidt norm of the matrix associated to A , as in Subsection 7.2.3. This defines a norm on each of $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ and $\mathcal{L}(\mathbf{C}^n, \mathbf{C}^m)$, which may be called the *Hilbert–Schmidt norm* as well.

7.5 Operator norms

Let m and n be positive integers, and let A be a linear mapping from \mathbf{R}^n into \mathbf{R}^m , or from \mathbf{C}^n into \mathbf{C}^m . Also let N_0 be a norm on \mathbf{R}^n or \mathbf{C}^n , and let N be a norm on \mathbf{R}^m or \mathbf{C}^m , as appropriate. The corresponding *operator norm*

$$(7.5.1) \quad \|A\|_{op} = \|A\|_{op, N_0 N}$$

of A is defined to be the supremum of

$$(7.5.2) \quad \frac{N(A(v))}{N_0(v)}$$

over all $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, with $v \neq 0$. Note that this ratio is bounded, by (7.3.8), so that the supremum is finite.

More precisely, (7.3.8) is equivalent to saying that the nonnegative real number C is an upper bound for (7.5.2) when $v \neq 0$, because (7.3.8) holds automatically when $v = 0$. Thus (7.3.8) holds exactly when

$$(7.5.3) \quad \|A\|_{op} \leq C.$$

In particular, (7.3.8) holds with $C = \|A\|_{op}$. Alternatively, $\|A\|_{op}$ is the infimum of the nonnegative real numbers C for which (7.3.8) holds.

7.5.1 Two other characterizations of the operator norm

Note that

$$(7.5.4) \quad \frac{N(A(v))}{N_0(v)} = N((1/N_0(v)) A(v)) = N(A((1/N_0(v)) v))$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, with $v \neq 0$. Similarly,

$$(7.5.5) \quad N_0((1/N_0(v)) v) = N_0(v)/N_0(v) = 1$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, with $v \neq 0$. Using this, one can verify that $\|A\|_{op}$ is the same as the supremum of

$$(7.5.6) \quad N(A(v))$$

over all $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, such that $N_0(v) = 1$. In both cases, one takes the supremum of the same set of nonnegative real numbers, although this set is described in slightly different ways.

This is also the same as the supremum of (7.5.6) over all $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, with $N_0(v) \leq 1$. Of course, (7.5.6) is equal to 0 when $v = 0$, and otherwise

$$(7.5.7) \quad N(A(v)) = N(A((1/N_0(v))v)) N_0(v)$$

when $v \neq 0$, as in (7.5.4). If $N_0(v) \leq 1$, then it follows that

$$(7.5.8) \quad N(A(v)) \leq N(A((1/N_0(v))v)).$$

This implies that the supremum of (7.5.6) over $v \in \mathbf{R}^n$ with $N_0(v) \leq 1$ is the same as the supremum of (7.5.6) over $v \in \mathbf{R}^n$ with $N_0(v) = 1$, because of (7.5.5).

7.5.2 Checking that the operator norm is a norm

By construction, $\|A\|_{op}$ is a nonnegative real number, and

$$(7.5.9) \quad \|A\|_{op} = 0$$

if and only if $A(v) = 0$ for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate.

If $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then tA also defines a linear mapping from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m , as appropriate. It is easy to see that

$$(7.5.10) \quad \|tA\|_{op} = |t| \|A\|_{op},$$

because

$$(7.5.11) \quad N(tA(v)) = |t| N(A(v))$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate.

Let B be another linear mapping from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m , as appropriate, so that $A + B$ defines another such linear mapping. Observe that

$$(7.5.12) \quad \begin{aligned} N((A+B)(v)) &= N(A(v) + B(v)) &\leq N(A(v)) + N(B(v)) \\ &\leq \|A\|_{op} N_0(v) + \|B\|_{op} N_0(v) \\ &= (\|A\|_{op} + \|B\|_{op}) N_0(v) \end{aligned}$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. This implies that

$$(7.5.13) \quad \|A + B\|_{op} \leq \|A\|_{op} + \|B\|_{op}.$$

Thus the operator norm satisfies the requirements of a norm on the space of linear mappings from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m , as appropriate.

7.5.3 Comparing this with the Hilbert–Schmidt norm

Suppose for the moment that N_0 is the standard Euclidean norm on \mathbf{R}^n or \mathbf{C}^n , and that N is the standard Euclidean norm on \mathbf{R}^m or \mathbf{C}^m , as appropriate. In this case, we get that

$$(7.5.14) \quad \|A\|_{op} \leq \|A\|_{HS},$$

as in (7.2.10), where the right side is the Hilbert–Schmidt norm of A , as in the previous section. One can check that

$$(7.5.15) \quad \|A\|_{HS} \leq n^{1/2} \|A\|_{op},$$

directly from the definitions of these two norms.

7.5.4 The $n = m$ case

Suppose now that $n = m$, and that $N_0 = N$ on \mathbf{R}^n or \mathbf{C}^n , as appropriate. Let I be the *identity operator* on \mathbf{R}^n or \mathbf{C}^n , as appropriate, so that $I(v) = v$ for all $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. It is easy to see that

$$(7.5.16) \quad \|I\|_{op} = 1$$

under these conditions.

We also have that

$$(7.5.17) \quad \|I\|_{HS} = n^{1/2}.$$

This corresponds to taking $A = I$ in (7.5.22).

Let A be a linear mapping from \mathbf{R}^n or \mathbf{C}^n into itself, as appropriate, and let $[a_{j,l}]$ be the corresponding $n \times n$ matrix or real or complex numbers, as in Section 7.2. Suppose that $[a_{j,l}]$ is a diagonal matrix, so that $a_{j,l} = 0$ unless $j = l$. This means that

$$(7.5.18) \quad (A(v))_j = a_{j,j} v_j$$

for each $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, and $j = 1, \dots, n$, as in (7.2.3). Note that the identity operator I corresponds to the *identity matrix*, the diagonal matrix whose diagonal entries are all equal to 1.

If $N_0 = N$ is the standard Euclidean norm on \mathbf{R}^n or \mathbf{C}^n , as appropriate, then it is not difficult to show that

$$(7.5.19) \quad \|A\|_{op} = \max_{1 \leq j \leq n} |a_{j,j}|$$

in this case. More precisely, one can first check that

$$(7.5.20) \quad \|A\|_{op} \leq \max_{1 \leq j \leq n} |a_{j,j}|,$$

directly from the definitions. To get the opposite inequality, one can use the fact that

$$(7.5.21) \quad A(e_k) = a_{k,k} e_k$$

for each $k = 1, \dots, n$, where e_k is the k th standard basis vector for \mathbf{R}^n or \mathbf{C}^n , as in Subsection 7.2.2. This also works when $N_0 = N$ is either of the norms $\|\cdot\|_1$ in (7.3.4) or $\|\cdot\|_\infty$ in (7.3.6).

Observe that

$$(7.5.22) \quad \|A\|_{HS} = \left(\sum_{j=1}^n |a_{j,j}|^2 \right)^{1/2}$$

in this situation.

7.6 Determinants and volumes

Let n be a positive integer, and let $[a_{j,l}]$ be an $n \times n$ matrix of real or complex numbers. The *determinant*

$$(7.6.1) \quad \det[a_{j,l}]$$

can be defined as a real or complex number, as appropriate, in a standard way. If $[a_{j,l}]$ is a diagonal matrix, so that $a_{j,l} = 0$ unless $j = l$, then it is well known that

$$(7.6.2) \quad \det[a_{j,l}] = \prod_{k=1}^n a_{k,k}.$$

Let A be a linear mapping from \mathbf{R}^n or \mathbf{C}^n into itself, and let $[a_{j,l}]$ be the corresponding $n \times n$ matrix of real or complex numbers, as appropriate, as in Section 7.2. The *determinant* of A is defined by

$$(7.6.3) \quad \det A = \det[a_{j,l}].$$

The determinant of the identity mapping I on \mathbf{R}^n or \mathbf{C}^n is equal to 1. If t is a real or complex number, as appropriate, then it is well known that

$$(7.6.4) \quad \det(tA) = t^n \det A.$$

7.6.1 Linear mappings and volumes

Let A be a linear mapping from \mathbf{R}^n into itself, and let E be a subset of \mathbf{R}^n . It is well known that

$$(7.6.5) \quad \text{Vol}_n(A(E)) = |\det A| \text{Vol}_n(E),$$

where $\text{Vol}_n(\cdot)$ is the usual n -dimensional volume of a subset of \mathbf{R}^n . This is discussed in many textbooks, and one may also be interested in the articles [41, 115].

More precisely, E should be sufficiently nice for the volume to be defined, depending on the definition of the volume being used. The right side of (7.6.5) should be interpreted as being equal to 0 when $\det A = 0$, even if $\text{Vol}_n(E)$ may be $+\infty$. Note that $|\det A|$ is uniquely determined by (7.6.5) when $\text{Vol}_n(E)$ is positive and finite.

If $t \in \mathbf{R}$, then put

$$(7.6.6) \quad tE = \{tv : v \in E\}.$$

It is well known that

$$(7.6.7) \quad \text{Vol}_n(tE) = |t|^n \text{Vol}_n(E),$$

where the right side is interpreted as being equal to 0 when $t = 0$, even if $\text{Vol}_n(E) = +\infty$. This corresponds to (7.6.5), with $A = tI$. See Section A.9 for more on the n -dimensional volume of subsets of \mathbf{R}^n .

7.6.2 Determinants and operator norms

Let N be a norm on \mathbf{R}^n , and let

$$(7.6.8) \quad \|A\|_{op} = \|A\|_{op, NN}$$

be the corresponding operator norm of A , as in the previous section. More precisely, this uses N as the norm on \mathbf{R}^n as both the domain and the range of A . It is well known that

$$(7.6.9) \quad |\det A| \leq \|A\|_{op}^n$$

under these conditions.

Suppose for the moment that N is the standard Euclidean norm on \mathbf{R}^n . If A can be diagonalized using an orthonormal basis for \mathbf{R}^n with respect to the standard inner product, then (7.6.9) can be verified directly. Otherwise, one can reduce to that case, by considering the composition of A with its adjoint with respect to the standard inner product on \mathbf{R}^n .

Similarly, suppose for the moment again that A corresponds to a diagonal matrix $[a_{j,i}]$ in the usual way, so that the determinant of A is as in (7.6.2). If N is one of the usual norms $\|\cdot\|_1$, $\|\cdot\|_2$, or $\|\cdot\|_\infty$ on \mathbf{R}^n , as in Subsection 7.3.1, then (7.6.9) may be obtained from (7.5.19).

7.6.3 Volumes and operator norms

If N is any norm on \mathbf{R}^n , then let \overline{B}_N be the closed unit ball in \mathbf{R}^n with respect to N , as in Subsection A.6.1. Thus, if $v \in \overline{B}_N$, then $N(v) \leq 1$, so that

$$(7.6.10) \quad N(A(v)) \leq \|A\|_{op}.$$

This means that

$$(7.6.11) \quad A(\overline{B}_N) \subseteq \|A\|_{op} \overline{B}_N,$$

where the right side is as in (7.6.6). In fact, $\|A\|_{op} \overline{B}_N$ is the same as the closed ball in \mathbf{R}^n centered at 0 with radius $\|A\|_{op}$ with respect to the metric $d_N(v, w) = N(v - w)$ associated to N .

Using (7.6.11), we get that

$$(7.6.12) \quad \text{Vol}_n(A(\overline{B}_N)) \leq \text{Vol}_n(\|A\|_{op} \overline{B}_N) = \|A\|_{op}^n \text{Vol}_n(\overline{B}_N),$$

where the second step is as in (7.6.7). We also have that

$$(7.6.13) \quad \text{Vol}_n(A(\overline{B}_N)) = |\det A| \text{Vol}_n(\overline{B}_N),$$

as in (7.6.5). One can use this to get (7.6.9), because $\text{Vol}_n(\overline{B}_N)$ is positive and finite. This uses the fact that N and the standard Euclidean norm on \mathbf{R}^n are each bounded by constant multiples of the other, as in Sections 6.10 and 6.11.

7.6.4 Eigenvalues and operator norms

Let A be a linear mapping from \mathbf{R}^n or \mathbf{C}^n into itself again. If $v \in \mathbf{R}^n$ or \mathbf{C}^n satisfies

$$(7.6.14) \quad A(v) = \lambda v$$

for some $\lambda \in \mathbf{R}$ or \mathbf{C} , as appropriate, then v is said to be an *eigenvector* of A with *eigenvalue* λ . More precisely, one may ask that (7.6.14) hold with $v \neq 0$ in order for λ to be considered as an eigenvalue of A .

Let N be a norm on \mathbf{R}^n or \mathbf{C}^n , as appropriate, and let $\|A\|_{op}$ be the corresponding operator norm of A , using N on both the domain and the range. Observe that

$$(7.6.15) \quad |\lambda| N(v) = N(\lambda v) = N(A(v)) \leq \|A\|_{op} N(v),$$

by (7.6.14) and the definition of the operator norm. This implies that

$$(7.6.16) \quad |\lambda| \leq \|A\|_{op}$$

when $v \neq 0$.

In the complex case, it is well known that the determinant of A is equal to the product of its n eigenvalues, associated to nonzero eigenvectors, counted with their appropriate multiplicities. One can use this to get (7.6.9) from (7.6.16).

7.7 Lipschitz constants

Let (X, d_X) and (Y, d_Y) be (nonempty) metric spaces. Remember that a mapping f from X into Y is said to be *Lipschitz* if

$$(7.7.1) \quad d_Y(f(x), f(w)) \leq C d_X(x, w)$$

for some nonnegative real number C and every $x, w \in X$. In this case, we may also say that f is Lipschitz with constant C , to make the role of C more explicit. Let $\text{Lip}(X, Y)$ be the space of all Lipschitz mappings from X into Y .

7.7.1 The definition of $\text{Lip}(f)$

Let f be a mapping from X into Y again, and note that (7.7.1) holds automatically when $x = w$. If x, w are distinct elements of X , then (7.7.1) is the same as saying that

$$(7.7.2) \quad \frac{d_Y(f(x), f(w))}{d_X(x, w)} \leq C.$$

If f is Lipschitz, and X has at least two elements, then put

$$(7.7.3) \quad \text{Lip}(f) = \text{Lip}_{X,Y}(f) = \sup \left\{ \frac{d_Y(f(x), f(w))}{d_X(x, w)} : x, w \in X, x \neq w \right\}.$$

Otherwise, if X has only one element, then we take $\text{Lip}(f) = 0$. Thus (7.7.1) holds for some $C \geq 0$ and every $x, w \in X$ if and only if f is Lipschitz and

$$(7.7.4) \quad \text{Lip}(f) \leq C.$$

In particular, if f is Lipschitz, then (7.7.1) holds with $C = \text{Lip}(f)$. Equivalently, $\text{Lip}(f)$ is the infimum of the nonnegative real numbers C such that (7.7.1) holds for every $x, w \in X$.

7.7.2 Operator norms and Lipschitz constants

Let m and n be positive integers, and suppose for the moment that $X = \mathbf{R}^n$ and $Y = \mathbf{R}^m$, or that $X = \mathbf{C}^n$ and $Y = \mathbf{C}^m$. Also let N_0 be a norm on \mathbf{R}^n or \mathbf{C}^n , and let N be a norm on \mathbf{R}^m or \mathbf{C}^m , as appropriate. Thus N_0 determines a metric d_{N_0} on \mathbf{R}^n or \mathbf{C}^n , and N determines a metric d_N on \mathbf{R}^m or \mathbf{C}^m , as appropriate.

Let A be a linear mapping from \mathbf{R}^n into \mathbf{R}^m , or from \mathbf{C}^n into \mathbf{C}^m , as appropriate. Under these conditions, A is Lipschitz with respect to d_{N_0} and d_N , as in (7.3.10). More precisely,

$$(7.7.5) \quad \text{Lip}(A) = \|A\|_{op},$$

where $\text{Lip}(A)$ is defined using the metrics d_{N_0} and d_N as in the preceding subsection, and the operator norm $\|A\|_{op}$ of A is defined using N_0 and N as in Section 7.5. This is easy to verify, directly from the definitions. This is similar to some of the remarks in Subsection 7.3.2.

7.7.3 Vector-valued Lipschitz functions

Let (X, d_X) be any nonempty metric space again, and let m be a positive integer. Let us take $Y = \mathbf{R}^m$ or \mathbf{C}^m , equipped with a norm N , which determines a metric d_N in the usual way. In this situation, if f is a mapping from X into Y , then (7.7.1) is the same as saying that

$$(7.7.6) \quad N(f(x) - f(w)) = d_N(f(x), f(w)) \leq C d_X(x, w).$$

Let t be a real or complex number, as appropriate, so that tf also defines a mapping from X into Y , and

$$(7.7.7) \quad \begin{aligned} d_N(tf(x), tf(w)) &= N(tf(x) - tf(w)) = N(t(f(x) - f(w))) \\ &= |t| N(f(x) - f(w)) = |t| d_N(f(x), f(w)) \end{aligned}$$

for every $x, w \in X$. If f is Lipschitz, then it is easy to see that tf is Lipschitz too, with

$$(7.7.8) \quad \text{Lip}(tf) = |t| \text{Lip}(f).$$

Similarly, let g be another mapping from X into Y , so that $f + g$ defines a mapping from X into Y as well. Observe that

$$\begin{aligned}
 d_N((f+g)(x), (f+g)(w)) &= N((f(x) + g(x)) - (f(w) + g(w))) \\
 (7.7.9) \qquad \qquad \qquad &= N((f(x) - f(w)) + (g(x) - g(w))) \\
 &\leq N(f(x) - f(w)) + N(g(x) - g(w))
 \end{aligned}$$

for every $x, w \in X$. If f and g are both Lipschitz, then it follows that

$$\begin{aligned}
 d_N((f+g)(x), (f+g)(w)) &\leq \text{Lip}(f) d_X(x, w) + \text{Lip}(g) d_X(x, w) \\
 (7.7.10) \qquad \qquad \qquad &= (\text{Lip}(f) + \text{Lip}(g)) d_X(x, w)
 \end{aligned}$$

for every $x, w \in X$. This implies that $f + g$ is Lipschitz, with

$$(7.7.11) \qquad \qquad \text{Lip}(f + g) \leq \text{Lip}(f) + \text{Lip}(g).$$

In particular, the space of Lipschitz mappings from X into \mathbf{R}^m or \mathbf{C}^m may be considered as a vector space over the real or complex numbers, as appropriate, with respect to pointwise addition and scalar multiplication. In the terminology of Subsection A.6.1, $\text{Lip}(f)$ defines a seminorm on this vector space. More precisely, $\text{Lip}(f) = 0$ if and only if f is a constant mapping on X .

7.8 Compositions and isometries

Let (X, d_X) , (Y, d_Y) , and (Z, d_Z) be (nonempty) metric spaces. If f is a mapping from X into Y , and g is a mapping from Y into Z , then the composition $g \circ f$ can be defined as a mapping from X into Z , with $(g \circ f)(x) = g(f(x))$ for every $x \in X$, as usual.

Suppose that f and g are both Lipschitz, so that

$$\begin{aligned}
 (7.8.1) \qquad d_Z((g \circ f)(x), (g \circ f)(w)) &= d_Z(g(f(x)), g(f(w))) \\
 &\leq \text{Lip}(g) d_Y(f(x), f(w)) \\
 &\leq \text{Lip}(f) \text{Lip}(g) d_X(x, w)
 \end{aligned}$$

for every $x, w \in X$. This implies that $g \circ f$ is Lipschitz as well, with

$$(7.8.2) \qquad \qquad \text{Lip}(g \circ f) \leq \text{Lip}(f) \text{Lip}(g).$$

More precisely, $\text{Lip}(f) = \text{Lip}_{X,Y}(f)$, $\text{Lip}(g) = \text{Lip}_{Y,Z}(g)$, and $\text{Lip}(g \circ f) = \text{Lip}_{X,Z}(g \circ f)$ are as defined in the previous section, using the appropriate metric spaces in the domains and ranges of these mappings.

7.8.1 Isometric mappings

A mapping f from X into Y is said to be an *isometry* if

$$(7.8.3) \qquad \qquad d_Y(f(x), f(w)) = d_X(x, w)$$

for every $x, w \in X$. In particular, this implies that f is Lipschitz, with constant $C = 1$. If X has at least two elements, then (7.8.3) implies that $\text{Lip}(f) = 1$.

If a mapping g from Y into Z is an isometry too, then it is easy to see that the composition $g \circ f$ of f and g is an isometric mapping from X into Z . This is basically the same as (7.8.1), with equality in the second and third steps.

Note that an isometric mapping f from X into Y is automatically injective. If f maps X onto Y , then the corresponding inverse mapping f^{-1} is an isometry from Y onto X . Otherwise, f^{-1} may be considered as an isometry from $f(X)$ onto X , using the restriction of d_Y to $f(X)$.

7.8.2 Composing linear mappings

Let n, m , and k be positive integers. Suppose that either A is a linear mapping from \mathbf{R}^n into \mathbf{R}^m and B is a linear mapping from \mathbf{R}^m into \mathbf{R}^k , or that A is a linear mapping from \mathbf{C}^n into \mathbf{C}^m and B is a linear mapping from \mathbf{C}^m into \mathbf{C}^k . Thus the composition $B \circ A$ is either defined as a mapping from \mathbf{R}^n into \mathbf{R}^k , or as a mapping from \mathbf{C}^n into \mathbf{C}^k .

In both cases, it is easy to see that $B \circ A$ is a linear mapping. Note that the matrix associated to $B \circ A$ can be given in terms of the matrices associated to A and B using matrix multiplication.

7.8.3 Compositions and operator norms

Let N_1, N_2 , and N_3 be norms on $\mathbf{R}^n, \mathbf{R}^m$, and \mathbf{R}^k , respectively, or on $\mathbf{C}^n, \mathbf{C}^m$, and \mathbf{C}^k , respectively, as appropriate. Using these norms, the operator norms $\|A\|_{op} = \|A\|_{op, N_1 N_2}$, $\|B\|_{op} = \|B\|_{op, N_2 N_3}$, and $\|B \circ A\|_{op} = \|B \circ A\|_{op, N_1 N_3}$ can be defined as in Section 7.5.

If $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, then

$$\begin{aligned} (7.8.4) \quad N_3((B \circ A)(v)) = N_3(B(A(v))) &\leq \|B\|_{op} N_2(A(v)) \\ &\leq \|A\|_{op} \|B\|_{op} N_1(v). \end{aligned}$$

It follows that

$$(7.8.5) \quad \|B \circ A\|_{op} \leq \|A\|_{op} \|B\|_{op}.$$

This could also be obtained from (7.8.2), using (7.7.5).

7.8.4 Isometric linear mappings

In this situation, A is said to be an *isometric linear mapping* with respect to N_1 and N_2 if

$$(7.8.6) \quad N_2(A(v)) = N_1(v)$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. This implies that

$$\begin{aligned} (7.8.7) \quad d_{N_2}(A(v), A(w)) = N_2(A(v) - A(w)) &= N_2(A(v - w)) \\ &= N_1(v - w) = d_{N_1}(v, w) \end{aligned}$$

for every $v, w \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, where d_{N_1} and d_{N_2} are the metrics associated to N_1 and N_2 , respectively. Conversely, (7.8.6) follows from (7.8.7) by taking $w = 0$. Thus A is isometric as a linear mapping with respect to N_1 and N_2 if and only if A is isometric with respect to the corresponding metrics d_{N_1} and d_{N_2} .

If A is an isometric linear mapping, then $\|A\|_{op} = \|A\|_{op, N_1 N_2} = 1$ in particular. If B is also an isometric linear mapping with respect to N_2 and N_3 , then their composition $B \circ A$ is an isometric linear mapping with respect to N_1 and N_3 . This is essentially the same as in (7.8.4), with equality in the second and third steps.

7.9 Bilipschitz embeddings

Let (X, d_X) and (Y, d_Y) be (nonempty) metric spaces, and let f be a mapping from X into Y . Also let c be a positive real number, and consider the following condition: for every $x, w \in X$, we have that

$$(7.9.1) \quad c d_X(x, w) \leq d_Y(f(x), f(w)).$$

In particular, this condition implies that f is injective.

If f is a one-to-one mapping from X onto Y , then (7.9.1) is the same as saying that the corresponding inverse mapping f^{-1} satisfies

$$(7.9.2) \quad d_X(f^{-1}(y), f^{-1}(z)) \leq (1/c) d_Y(y, z)$$

for every $y, z \in Y$. This means that f^{-1} is Lipschitz with constant $1/c$ as a mapping from Y into X .

Otherwise, if f is injective but not necessarily surjective, then one can consider the inverse mapping f^{-1} as a mapping from the image $f(X)$ of X under f into X . One may also consider $f(X)$ as a metric space, using the restriction of d_Y to $f(X)$.

7.9.1 Bilipschitz mappings

A mapping f from X into Y is said to be *bilipschitz* if f is Lipschitz and (7.9.1) holds for some $c > 0$. To be more precise, one may say that f is bilipschitz with constant $C \geq 1$ if f is Lipschitz with constant C and (7.9.1) holds with $c = 1/C$. Using this terminology, an isometric mapping from X into Y is the same as a bilipschitz mapping with constant $C = 1$.

A one-to-one mapping f from X onto Y is bilipschitz if and only if f is Lipschitz and the inverse mapping f^{-1} is Lipschitz as a mapping from Y into X . If f is not surjective, then one can consider the inverse mapping f^{-1} as a mapping from $f(X)$ into X , using the restriction of d_Y to $f(X)$, as before.

7.9.2 Compositions of bilipschitz mappings

Let (Z, d_Z) be another metric space, and let g be a mapping from Y into Z . Suppose that there is a positive real number c' such that

$$(7.9.3) \quad c' d_Y(y, u) \leq d_Z(g(y), g(u))$$

for every $u, y \in Y$. If $f : X \rightarrow Y$ satisfies (7.9.1) for some $c > 0$, then the composition $g \circ f$ satisfies an analogous condition as a mapping from X into Z . More precisely, for each $x, w \in X$, we have that

$$(7.9.4) \quad \begin{aligned} c c' d_X(x, w) \leq c' d_Y(f(x), f(w)) &\leq d_Z(g(f(x)), g(f(w))) \\ &= d_Z((g \circ f)(x), (g \circ f)(w)). \end{aligned}$$

In particular, if f and g are both bilipschitz, then $g \circ f$ is bilipschitz as a mapping from X into Z .

7.9.3 A condition on linear mappings

Let m and n be positive integers, and let A be a linear mapping from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m . Also let N_0 and N be norms on \mathbf{R}^n and \mathbf{R}^m , respectively, or on \mathbf{C}^n and \mathbf{C}^m , respectively, as appropriate. Suppose that

$$(7.9.5) \quad c N_0(v) \leq N(A(v))$$

for some positive real number c and every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. If $v, w \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, then we get that

$$(7.9.6) \quad \begin{aligned} c d_{N_0}(v, w) = c N_0(v - w) &\leq N(A(v - w)) \\ &= N(A(v) - A(w)) = d_N(A(v), A(w)), \end{aligned}$$

where d_{N_0} and d_N are the metrics associated to N_0 and N , respectively. Of course, (7.9.6) corresponds to (7.9.1) in this situation.

Note that (7.9.6) implies (7.9.5), by taking $w = 0$. Observe too that (7.9.5) implies directly that $v = 0$ when $A(v) = 0$. This is the same as saying that the kernel of A is trivial.

7.9.4 More on composing linear mappings

Let k be another positive integer, and let B be a linear mapping from \mathbf{R}^m into \mathbf{R}^k or from \mathbf{C}^m into \mathbf{C}^k , as appropriate. Also let N_3 be a norm on \mathbf{R}^k or \mathbf{C}^k , as appropriate. Suppose that

$$(7.9.7) \quad c' N(y) \leq N_3(B(y))$$

for some positive real number c' and every $y \in \mathbf{R}^m$ or \mathbf{C}^m , as appropriate. If $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, then we get that

$$(7.9.8) \quad c c' N_0(v) \leq c' N(A(v)) \leq N_3(B(A(v))) = N_3((B \circ A)(v)).$$

This basically corresponds to (7.9.4) in this situation, as before.

7.10 Linear mappings and seminorms

Let m and n be positive integers, and let A be a linear mapping from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m . Also let N be a seminorm on \mathbf{R}^m or \mathbf{C}^m , as appropriate, as in Subsection A.6.1. More precisely, the definition of a seminorm is given there in the real case, and the complex case is analogous.

Under these conditions, one can check that

$$(7.10.1) \quad N(A(v))$$

defines a seminorm on \mathbf{R}^n or \mathbf{C}^n , as appropriate. Indeed, if $v \in \mathbf{R}^n$ or \mathbf{C}^n and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then

$$(7.10.2) \quad N(A(tv)) = N(tA(v)) = |t|N(A(v)).$$

Similarly, if $w \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, too, then

$$(7.10.3) \quad N(A(v+w)) = N(A(v) + A(w)) \leq N(A(v)) + N(A(w)).$$

7.10.1 Using a norm N

Suppose for the rest of the section that N is a norm on \mathbf{R}^m or \mathbf{C}^m , as appropriate, so that

$$(7.10.4) \quad N(A(v)) = 0$$

only when $A(v) = 0$. If the kernel of A is trivial, then (7.10.4) holds only when $v = 0$. This means that

$$(7.10.5) \quad N(A(v)) \text{ defines a norm on } \mathbf{R}^n \text{ or } \mathbf{C}^n,$$

as appropriate.

Let N_0 be another norm on \mathbf{R}^n or \mathbf{C}^n , as appropriate. Under these conditions, one can get that (7.9.5) holds for some $c > 0$, using (7.10.5) and the remarks in Subsections 6.10.3 and 6.11.3.

7.10.2 The $m = n$ case

Suppose now that $m = n$, so that A is a linear mapping from \mathbf{R}^n into itself, or from \mathbf{C}^n into itself. If the kernel of A is trivial, then it is well known that A maps \mathbf{R}^n or \mathbf{C}^n onto itself, as appropriate. This means that A has an inverse mapping A^{-1} on \mathbf{R}^n or \mathbf{C}^n , as appropriate. It is easy to see that the inverse mapping is also linear under these conditions.

In this case, (7.9.5) can be reformulated as saying that

$$(7.10.6) \quad c N_0(A^{-1}(u)) \leq N(u)$$

for every $u \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. Equivalently, this means that

$$(7.10.7) \quad N_0(A^{-1}(u)) \leq (1/c) N(u)$$

for every $u \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate.

We have seen previously that this type of condition holds, because A^{-1} is a linear mapping on \mathbf{R}^n or \mathbf{C}^n , as appropriate. More precisely, this condition holds with

$$(7.10.8) \quad 1/c = \|A^{-1}\|_{op} = \|A^{-1}\|_{op, NN_0},$$

where the operator norm of A^{-1} is defined using N on the domain of A^{-1} , and N_0 on the range of A^{-1} .

Note that $\|A^{-1}\|_{op} > 0$, because $A^{-1} \neq 0$. Of course, if (7.10.7) holds for some $c > 0$, then we have that

$$(7.10.9) \quad \|A^{-1}\|_{op} \leq 1/c.$$

7.11 Small perturbations

Let (X, d_X) be a nonempty metric space, let m be a positive integer, and let f be a mapping from X into \mathbf{R}^m or \mathbf{C}^m . Also let N be a norm on \mathbf{R}^m or \mathbf{C}^m , as appropriate, which leads to a metric d_N in the usual way.

Suppose that

$$(7.11.1) \quad c d_X(x, w) \leq d_N(f(x), f(w)) = N(f(x) - f(w))$$

for some positive real number c and all $x, w \in X$. Let g be another mapping from X into \mathbf{R}^m or \mathbf{C}^m , as appropriate. We would like to show that g satisfies an analogous condition when g is sufficiently close to f , in a suitable sense.

More precisely, we ask first that $f - g$ be Lipschitz with respect to d_N on the range. This means that

$$(7.11.2) \quad \begin{aligned} N((f(x) - g(x)) - (f(w) - g(w))) &= d_N(f(x) - g(x), f(w) - g(w)) \\ &\leq \text{Lip}(f - g) d_X(x, w) \end{aligned}$$

for every $x, w \in X$, where $\text{Lip}(f - g)$ is as defined in Subsection 7.7.1.

Of course,

$$(7.11.3) \quad \begin{aligned} N(f(x) - f(w)) &\leq N((f(x) - g(x)) - (f(w) - g(w))) \\ &\quad + N(g(x) - g(w)) \end{aligned}$$

for every $x, w \in X$, by the triangle inequality. Combining this with (7.11.1) and (7.11.2), we get that

$$(7.11.4) \quad c d_X(x, w) \leq N(g(x) - g(w)) + \text{Lip}(f - g) d_X(x, w)$$

for every $x, w \in X$.

It follows that

$$(7.11.5) \quad (c - \text{Lip}(f - g)) d_X(x, w) \leq N(g(x) - g(w)) = d_N(g(x), g(w))$$

for every $x, w \in X$. If $\text{Lip}(f - g) < c$, then this is the same type of condition as before.

7.11.1 Analogous arguments for linear mappings

Let n be a positive integer, and let us now take $X = \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. Let N_0 be a norm on \mathbf{R}^n or \mathbf{C}^n , as appropriate, which leads to a metric d_{N_0} , as usual. Also let A be a linear mapping from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m , as appropriate.

Suppose that

$$(7.11.6) \quad c N_0(v) \leq N(A(v))$$

for some $c > 0$ and all $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. Let B be another linear mapping from \mathbf{R}^n into \mathbf{R}^m or from \mathbf{C}^n into \mathbf{C}^m , as appropriate. We would like to verify that B satisfies the same type of condition when B is sufficiently close to A . This could be obtained from the remarks at the beginning of the section, but the analogous argument is a bit simpler in this case, as follows.

Remember that

$$(7.11.7) \quad N(A(v) - B(v)) \leq \|A - B\|_{op} N_0(v)$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, where the operator norm $\|A - B\|_{op}$ of $A - B$ is as defined in Section 7.5. Thus

$$(7.11.8) \quad \begin{aligned} N(A(v)) &\leq N(A(v) - B(v)) + N(B(v)) \\ &\leq \|A - B\|_{op} N_0(v) + N(B(v)) \end{aligned}$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, using the triangle inequality in the first step.

This implies that

$$(7.11.9) \quad c N_0(v) \leq N(B(v)) + \|A - B\|_{op} N_0(v)$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate, by (7.11.6). Hence

$$(7.11.10) \quad (c - \|A - B\|_{op}) N_0(v) \leq N(B(v))$$

for every $v \in \mathbf{R}^n$ or \mathbf{C}^n , as appropriate. This is the same type of condition as before when $\|A - B\|_{op} < c$.

In particular, if $\|A - B\|_{op} < c$, then (7.11.10) implies that the kernel of B is trivial, as mentioned in Subsection 7.9.3. If $m = n$, then it follows that B is invertible, as mentioned in Subsection 7.10.2.

7.12 The contraction mapping theorem

Let X be a set, and let f be a mapping from X into itself. We say that $x \in X$ is a *fixed point* of f on X if

$$(7.12.1) \quad f(x) = x.$$

Suppose now that $(X, d(x, y))$ is a nonempty metric space, and that f is Lipschitz mapping from X into itself with constant $c \geq 0$, so that

$$(7.12.2) \quad d(f(x), f(y)) \leq c d(x, y)$$

for every $x, y \in X$. If $c < 1$, and if X is complete with respect to $d(\cdot, \cdot)$, then a famous theorem states that f has a unique fixed point in X .

To get uniqueness, suppose that $x, y \in X$ are fixed points of f , so that

$$(7.12.3) \quad d(x, y) = d(f(x), f(y)) \leq c d(x, y),$$

by (7.12.2). If $c < 1$, then this implies that $d(x, y) = 0$, so that $x = y$. Note that this does not need completeness of X .

7.12.1 Existence of the fixed point

To get the existence of the fixed point, let x_0 be any element of X , and let x_1, x_2, x_3, \dots be the sequence of elements of X defined recursively by

$$(7.12.4) \quad x_j = f(x_{j-1})$$

when $j \geq 1$. Observe that

$$(7.12.5) \quad d(x_j, x_{j+1}) = d(f(x_{j-1}), f(x_j)) \leq c d(x_{j-1}, x_j)$$

for every $j \geq 1$. This implies that

$$(7.12.6) \quad d(x_j, x_{j+1}) \leq c^j d(x_0, x_1)$$

for every $j \geq 1$.

If $l < n$ are positive integers, then it follows that

$$(7.12.7) \quad d(x_l, x_n) \leq \sum_{j=l}^{n-1} d(x_j, x_{j+1}) \leq \sum_{j=l}^{n-1} c^j d(x_0, x_1),$$

using the triangle inequality in the first step. This is the same as saying that

$$(7.12.8) \quad d(x_l, x_n) \leq c^l \sum_{j=0}^{n-l-1} c^j d(x_0, x_1).$$

Remember that $\sum_{l=0}^{\infty} c^l$ is a convergent geometric series when $c < 1$, with sum equal to $1/(1-c)$. Using this, we get that

$$(7.12.9) \quad d(x_l, x_n) \leq \frac{c^l}{1-c} d(x_0, x_1).$$

This means that $\{x_j\}_{j=0}^{\infty}$ is a Cauchy sequence in X , because $c^l \rightarrow 0$ as $l \rightarrow \infty$ when $c < 1$.

If X is complete, then it follows that $\{x_j\}_{j=0}^{\infty}$ converges to an element x of X . We also have that

$$(7.12.10) \quad \{f(x_j)\}_{j=0}^{\infty} \text{ converges to } f(x) \text{ in } X,$$

because f is continuous. By construction, $\{f(x_j)\}_{j=0}^{\infty}$ is the same as $\{x_{j+1}\}_{j=0}^{\infty}$, which converges to x . Hence $f(x) = x$, by the uniqueness of the limit of a convergent sequence in a metric space.

7.12.2 Brouwer's fixed-point theorem

Let n be a positive integer, and let $\overline{B}(0, 1)$ be the closed unit ball in \mathbf{R}^n with respect to the standard Euclidean metric. Also let f be a continuous mapping from $\overline{B}(0, 1)$ into itself, with respect to the restriction of the standard Euclidean metric on \mathbf{R}^n to $\overline{B}(0, 1)$. Under these conditions, *Brouwer's fixed-point theorem* states that f has a fixed point in $\overline{B}(0, 1)$.

If $n = 1$, then this can be obtained from the intermediate value theorem. These and related matters are discussed in many textbooks, as well as the articles [32, 33, 36, 77, 150, 155, 174, 176, 182, 188, 203, 212, 224], for instance.

7.13 Norms on \mathbf{R}^n and completeness

Let n be a positive integer again, and let N be a norm on \mathbf{R}^n . Note that

(7.13.1) \mathbf{R}^n is complete with respect to the metric d_N associated to N .

This can be obtained from the completeness of \mathbf{R}^n with respect to the standard Euclidean metric, and the comparability of N with the standard Euclidean norm on \mathbf{R}^n , as in Subsections 6.10.3 and 6.11.3.

7.13.1 Small perturbations of the identity mapping on \mathbf{R}^n

Let g be a Lipschitz mapping from \mathbf{R}^n into itself, with respect to d_N , and with constant $c \geq 0$. This means that

$$(7.13.2) \quad N(g(x) - g(y)) \leq c N(x - y)$$

for every $x, y \in \mathbf{R}^n$ in this situation.

Let $a \in \mathbf{R}^n$ be given, and put

$$(7.13.3) \quad g_a(x) = g(x) + a$$

for every $x \in \mathbf{R}^n$. Observe that

$$(7.13.4) \quad g_a(x) - g_a(y) = g(x) - g(y)$$

for every $x, y \in \mathbf{R}^n$, so that

$$(7.13.5) \quad N(g_a(x) - g_a(y)) = N(g(x) - g(y)) \leq c N(x - y).$$

Thus g_a is also Lipschitz with constant c as a mapping from \mathbf{R}^n into itself, with respect to d_N .

Suppose that $c < 1$, so that the contraction mapping theorem can be applied to g_a on \mathbf{R}^n . It follows that there is a unique point $x(a) \in \mathbf{R}^n$ such that

$$(7.13.6) \quad g(x(a)) + a = g_a(x(a)) = x(a).$$

Put

$$(7.13.7) \quad h(x) = x - g(x)$$

for every $x \in \mathbf{R}^n$, which defines a mapping from \mathbf{R}^n into itself. The previous statement can be reformulated as saying that for every $a \in \mathbf{R}^n$ there is a unique $x(a) \in \mathbf{R}^n$ such that

$$(7.13.8) \quad h(x(a)) = x(a) - g(x(a)) = a.$$

Of course, this is the same as saying that h is a one-to-one mapping from \mathbf{R}^n onto itself.

Note that h is Lipschitz with constant $1 + c$ on \mathbf{R}^n with respect to d_N , as in Subsection 7.7.3. We also have that

$$(7.13.9) \quad (1 - c)N(x - w) \leq N(h(x) - h(w))$$

for every $x, w \in \mathbf{R}^n$, as in Section 7.11.

In the next section, we consider a version of this for mappings from a closed ball in \mathbf{R}^n into \mathbf{R}^n .

These and related results will be used in the proof of the inverse function theorem, as in Subsection 9.4.3.

7.14 A localized condition

Let n be a positive integer, and let N be a norm on \mathbf{R}^n again. Also let

$$(7.14.1) \quad \overline{B}_N(r) = \{x \in \mathbf{R}^n : N(x) \leq r\}$$

be the closed ball in \mathbf{R}^n centered at 0 with radius $r \geq 0$ with respect to the metric d_N on \mathbf{R}^n associated to N . Remember that this is a closed set in \mathbf{R}^n with respect to d_N for each $r \geq 0$, as in Subsection 1.9.3.

7.14.1 Completeness of $\overline{B}_N(r)$

It is easy to see that

$$(7.14.2) \quad \overline{B}_N(r) \text{ is complete as a metric space}$$

with respect to the restriction of d_N to $\overline{B}_N(r)$ for every $r \geq 0$, as in Subsection 1.7.1. More precisely, any Cauchy sequence of elements of $\overline{B}_N(r)$ with respect to d_N may be considered as a Cauchy sequence in \mathbf{R}^n with respect to d_N as well. Such a sequence converges to an element of \mathbf{R}^n with respect to d_N , because \mathbf{R}^n is complete with respect to d_N , as in Section 7.13. The limit of this sequence is contained in $\overline{B}_N(r)$, because $\overline{B}_N(r)$ is a closed set in \mathbf{R}^n with respect to d_N .

7.14.2 Lipschitz mappings on $\overline{B}_N(r)$

Let r be a positive real number, and let g be a Lipschitz mapping from $\overline{B}_N(r)$ into \mathbf{R}^n with constant $c \geq 0$, with respect to d_N and its restriction to $\overline{B}_N(r)$. Thus

$$(7.14.3) \quad N(g(x) - g(y)) \leq c N(x - y)$$

for every $x, y \in \overline{B}_N(r)$. Suppose also that

$$(7.14.4) \quad g(0) = 0.$$

This implies that

$$(7.14.5) \quad N(g(x)) \leq c N(x) \leq c r$$

for every $x \in \overline{B}_N(r)$, using (7.14.3) in the first step.

Suppose too that

$$(7.14.6) \quad c < 1,$$

and let $a \in \mathbf{R}^n$ be given, with

$$(7.14.7) \quad N(a) \leq (1 - c) r.$$

Put

$$(7.14.8) \quad g_a(x) = g(x) + a$$

for each $x \in \overline{B}_N(r)$, as before. Observe that

$$(7.14.9) \quad N(g_a(x)) \leq N(g(x)) + N(a) \leq c r + (1 - c) r = r$$

for every $x \in \overline{B}_N(r)$, using (7.14.5) and (7.14.7) in the second step.

7.14.3 Using the contraction mapping theorem on $\overline{B}_N(r)$

This shows that g_a maps $\overline{B}_N(r)$ into itself under these conditions. Note that g_a is Lipschitz with constant c with respect to d_N as well, as in (7.13.5). Hence we can apply the contraction mapping theorem to g_a on $\overline{B}_N(r)$, because of (7.14.6). This implies that there is a unique point $x(a) \in \overline{B}_N(r)$ such that

$$(7.14.10) \quad g(x(a)) + a = g_a(x(a)) = x(a).$$

Let h be the mapping from $\overline{B}_N(r)$ into \mathbf{R}^n defined by

$$(7.14.11) \quad h(x) = x - g(x)$$

for every $x \in \overline{B}_N(r)$. Observe that

$$(7.14.12) \quad h(x(a)) = x(a) - g(x(a)) = a,$$

by (7.14.10). It follows that $a \in h(\overline{B}_N(r))$, so that

$$(7.14.13) \quad h(\overline{B}_N(r)) \supseteq \overline{B}_N((1 - c) r)$$

in this situation.

Note that h is Lipschitz with constant $1 + c$ on $\overline{B}_N(r)$ with respect to d_N , and that (7.13.9) holds for all $x, w \in \overline{B}_N(r)$, as in Subsection 7.7.3 and Section 7.11.

7.15 Open mappings

Let (X, d_X) and (Y, d_Y) be metric spaces. A mapping f from X into Y is said to be an *open mapping* if for every open set $U \subseteq X$,

$$(7.15.1) \quad f(U) \text{ is an open set in } Y.$$

One can check that this happens if and only if for every $x \in X$ and $r > 0$ there is a $t > 0$ such that

$$(7.15.2) \quad B_Y(f(x), t) \subseteq f(B_X(x, r)).$$

7.15.1 A helpful criterion on \mathbf{R}^n

Let n be a positive integer, and let N be a norm on \mathbf{R}^n . If $v \in \mathbf{R}^n$ and $r > 0$, then let

$$(7.15.3) \quad B_N(v, r) = \{w \in \mathbf{R}^n : N(v - w) < r\}$$

and

$$(7.15.4) \quad \overline{B}_N(v, r) = \{w \in \mathbf{R}^n : N(v - w) \leq r\}$$

be the open and closed balls in \mathbf{R}^n centered at v with radius r with respect to the metric d_N associated to N .

Let $v_0 \in \mathbf{R}^n$ and $r_0 > 0$ be given, and let f be a mapping from $\overline{B}_N(v_0, r_0)$ into \mathbf{R}^n . Suppose that $f(x) - x$ is Lipschitz with constant $c \geq 0$ on $\overline{B}_N(v_0, r_0)$, with respect to d_N and its restriction to $\overline{B}_N(v_0, r_0)$. If $c < 1$, then

$$(7.15.5) \quad \overline{B}_N(f(v_0), (1 - c)r_0) \subseteq f(\overline{B}_N(v_0, r_0)).$$

This follows from the remarks in the previous section when $v_0 = f(v_0) = 0$. Otherwise, one can reduce to this case, using translations in \mathbf{R}^n .

7.15.2 Analogous statements for open balls

Similarly, let $v \in \mathbf{R}^n$ and $r > 0$ be given, and let f be a mapping from $B_N(v, r)$ into \mathbf{R}^n . Suppose that $f(x) - x$ is Lipschitz with constant c , $0 \leq c < 1$, on $B_N(v, r)$, with respect to d_N and its restriction to $B_N(v, r)$. Let $w \in B_N(v, r)$ be given, and let r_1 be a positive real number such that

$$(7.15.6) \quad r_1 \leq r - N(v - w).$$

Note that $r - N(v - w) > 0$, and that (7.15.6) implies that

$$(7.15.7) \quad B_N(w, r_1) \subseteq B_N(v, r),$$

by the triangle inequality.

If $0 < r_0 < r_1$, then $\overline{B}_N(w, r_0) \subseteq B_N(w, r_1)$, and

$$(7.15.8) \quad \overline{B}_N(f(w), (1 - c)r_0) \subseteq f(\overline{B}_N(w, r_0)),$$

as in (7.15.5). This implies that

$$(7.15.9) \quad B_N(f(w), (1 - c)r_1) \subseteq f(B_N(w, r_1)).$$

In particular, this means that f is an open mapping from $B_N(v, r)$ into \mathbf{R}^n , with respect to d_N and its restriction to $B_N(v, r)$.

Chapter 8

Some topics related to differentiability

8.1 An integral triangle inequality

Let m be a positive integer, and let N be a norm on \mathbf{R}^m or \mathbf{C}^m . If v_1, \dots, v_l are finitely many elements of \mathbf{R}^m or \mathbf{C}^m and t_1, \dots, t_l are real or complex numbers, as appropriate, then it is easy to see that

$$(8.1.1) \quad N\left(\sum_{k=1}^l t_k v_k\right) \leq \sum_{k=1}^l N(t_k v_k) = \sum_{k=1}^l |t_k| N(v_k).$$

We would like to consider analogous statements for integrals instead of finite sums. Although this works for Riemann–Stieltjes integrals, it is sufficient to consider Riemann integrals for the result in the next section.

8.1.1 Vector-valued functions on $[a, b]$

Let a and b be real numbers with $a < b$, and let f be a continuous function defined on the closed interval $[a, b]$ with values in \mathbf{R}^m or \mathbf{C}^m , as appropriate. This implicitly uses the restriction of the standard Euclidean metric on the real line to $[a, b]$, and one can also use the standard Euclidean metric on \mathbf{R}^m or \mathbf{C}^m , as appropriate. Equivalently, f can be given as

$$(8.1.2) \quad f(t) = (f_1(t), \dots, f_m(t)),$$

where f_1, \dots, f_m are continuous real or complex-valued functions on $[a, b]$, as appropriate. Of course, the continuity of a complex-valued function on $[a, b]$ is equivalent to the continuity of its real and imaginary parts, as real-valued functions on $[a, b]$.

8.1.2 Riemann–Stieltjes integrals of vector-valued functions

Let α be a monotonically increasing real-valued function on $[a, b]$. As before, it suffices to consider the case where $\alpha(t) = t$ for every $t \in [a, b]$ for the result discussed in the next section, which corresponds to using ordinary Riemann integrals on $[a, b]$. To define the Riemann–Stieltjes integral

$$(8.1.3) \quad \int_a^b f(t) d\alpha(t)$$

as an element of \mathbf{R}^m or \mathbf{C}^m , as appropriate, one can take the j th coordinate of (8.1.3) to be the Riemann–Stieltjes integral of f_j with respect to α on $[a, b]$ for each $j = 1, \dots, m$. Similarly, the Riemann–Stieltjes integral of a continuous complex-valued function on $[a, b]$ can be reduced to the real case, by considering the real and imaginary parts of the function.

8.1.3 Norms of integrals

Remember that N is (uniformly) continuous as a real-valued function on \mathbf{R}^m or \mathbf{C}^m , as appropriate, as in Subsection 6.11.1. Thus $N(f(t))$ defines a nonnegative real-valued continuous function on $[a, b]$. In particular, its Riemann–Stieltjes integral over $[a, b]$ with respect to α can be defined as a nonnegative real number in the usual way.

Under these conditions, one can check that

$$(8.1.4) \quad N\left(\int_a^b f(t) d\alpha(t)\right) \leq \int_a^b N(f(t)) d\alpha(t).$$

More precisely, these integrals can be approximated by Riemann–Stieltjes sums associated to sufficiently fine partitions of $[a, b]$. To get (8.1.4), one can use the analogous inequalities for Riemann–Stieltjes sums, which follow from (8.1.1). One can also consider Riemann–Stieltjes integrability conditions on $[a, b]$, instead of continuity.

If N is either of the norms $\|\cdot\|_1$ or $\|\cdot\|_\infty$ on \mathbf{R}^m , as in Subsection 1.3.2, then (8.1.4) can be obtained from standard properties of Riemann–Stieltjes integrals of real-valued functions on $[a, b]$. If N is the standard Euclidean norm on \mathbf{R}^m , then one can reduce to a standard property of Riemann–Stieltjes integrals of real-valued functions using the standard inner product on \mathbf{R}^m , as in the proof of (40) in Theorem 6.25 on p135 of [210]. There is an analogous argument that works for arbitrary norms, using additional results about norms that have not been discussed here.

Of course, complex-valued functions on $[a, b]$ may be considered as functions with values in \mathbf{R}^2 , and the absolute value of a complex number corresponds to the standard Euclidean norm on \mathbf{R}^2 .

8.2 A basic Lipschitz estimate

Let a and b be real numbers with $a < b$, and let m be a positive integer. Also let f be a function defined on $[a, b]$ with values in \mathbf{R}^m or \mathbf{C}^m .

The derivative $f'(t)$ of f at a point $t \in (a, b)$ can be defined in the usual way, as the limit of difference quotients, when the limit exists. Similarly, if $t = a$ or b , then one can consider the corresponding one-sided limit.

The differentiability of f at any $t \in [a, b]$ is equivalent to the differentiability of the j th component f_j of f at t for each $j = 1, \dots, m$, as a real or complex-valued function on $[a, b]$. Similarly, the differentiability of a complex-valued function on $[a, b]$ is equivalent to the differentiability of its real and imaginary parts. If f is differentiable at $t \in [a, b]$, then the j th coordinate of $f'(t)$ is equal to the derivative of f_j at t for each $j = 1, \dots, m$, and their common value is denoted $f'_j(t)$.

8.2.1 Continuously-differentiable vector-valued functions

Let us suppose from now on in this section that f is continuously differentiable on $[a, b]$. This means that the derivative $f'(t)$ exists for every $t \in [a, b]$, and that $f'(t)$ is continuous as a function on $[a, b]$ with values in \mathbf{R}^m or \mathbf{C}^m , as appropriate. Here we use the restriction of the standard Euclidean metric on \mathbf{R} to $[a, b]$, and the standard Euclidean metric on \mathbf{R}^m or \mathbf{C}^m , as appropriate.

Equivalently, f_j should be continuously differentiable as a real or complex-valued function on $[a, b]$ for each $j = 1, \dots, m$. Of course, $f'(t)$ is really a one-sided derivative when $t = a$ or b .

8.2.2 The fundamental theorem of calculus

If $x \in [a, b]$, then the fundamental theorem of calculus implies that

$$(8.2.1) \quad \int_a^x f'(t) dt = f(x) - f(a),$$

where the Riemann integral on the left side can be defined as in the previous section. Of course, this can be reduced to the case of real-valued functions in the usual way.

Similarly, if $a \leq w \leq x \leq b$, then

$$(8.2.2) \quad \int_w^x f'(t) dt = f(x) - f(w).$$

8.2.3 A Lipschitz condition

Let N be a norm on \mathbf{R}^m or \mathbf{C}^m , as appropriate. If $a \leq w \leq x \leq b$, then

$$(8.2.3) \quad N(f(x) - f(w)) = N\left(\int_w^x f'(t) dt\right) \leq \int_w^x N(f'(t)) dt.$$

This uses (8.2.2) in the first step, and (8.1.4) in the second step, applied to f' and $\alpha(t) = t$.

Note that $N(f'(t))$ is continuous as a real-valued function on $[a, b]$, as in the previous section. In particular, $N(f'(t))$ is bounded on $[a, b]$, because $[a, b]$ is compact, and in fact the maximum of $N(f'(t))$ is attained on $[a, b]$.

It follows from (8.2.3) that

$$(8.2.4) \quad N(f(x) - f(w)) \leq |x - w| \sup_{a \leq t \leq b} N(f'(t))$$

for every $x, w \in [a, b]$. Thus f is Lipschitz with respect to the restriction of the standard Euclidean metric on \mathbf{R} to $[a, b]$ and the metric d_N associated to N on \mathbf{R}^m or \mathbf{C}^m , as appropriate.

8.2.4 The Lipschitz constant of f

More precisely,

$$(8.2.5) \quad \text{Lip}(f) = \sup_{a \leq t \leq b} N(f'(t)),$$

where $\text{Lip}(f)$ is as defined in Subsection 7.7.1. The fact that $\text{Lip}(f)$ is less than or equal to the right side of (8.2.5) follows from (8.2.4).

To get the opposite inequality, it suffices to verify that $N(f'(t))$ is less than or equal to $\text{Lip}(f)$ for every $t \in [a, b]$. It is easy to see that that difference quotients used to define $f'(t)$ have norm less than or equal to $\text{Lip}(f)$, by definition of $\text{Lip}(f)$. This implies that $N(f'(t)) \leq \text{Lip}(f)$, by taking the limit of the difference quotients.

Remember that Lipschitz conditions like these for real-valued functions on $[a, b]$ may be obtained using the mean value theorem, as in Subsection 1.2.2. It is well known that the mean value theorem does not always work for vector-valued functions. However, one can get the same type of Lipschitz condition, as in (8.2.4). Theorem 5.9 on p113 of [210] is a version of this, using the standard Euclidean norm on \mathbf{R}^m . The proof uses the standard inner product on \mathbf{R}^m , to reduce to the case of real-valued functions on $[a, b]$.

This type of argument, reducing to the mean value theorem for real-valued functions, also works for functions that are continuous on $[a, b]$ and differentiable on (a, b) , with bounded derivative. If N is the norm $\|\cdot\|_\infty$ defined on \mathbf{R}^m as in Subsection 1.3.2, then it is not difficult to reduce directly to the case of real-valued functions. If N is the norm $\|\cdot\|_1$ defined on \mathbf{R}^m as before, then there is an argument analogous to the one in [210] for reducing to the case of real-valued functions. As in the previous section, there is an analogous argument that works for arbitrary norms, using results about norms that have not been discussed here. There are other variants of these types of arguments as well.

Complex-valued functions may be considered as functions with values in \mathbf{R}^2 , with the absolute value of a complex number corresponding to the standard Euclidean norm on \mathbf{R}^2 , as before.

8.3 Some partial Lipschitz conditions

Let n be a positive integer, and let E be a subset of \mathbf{R}^n . Also let (Y, d_Y) be a metric space, and let f be a mapping from E into Y . Suppose that l is a positive integer less than or equal to n , and that C_l is a nonnegative real number.

Let us say that f is *partially Lipschitz in the l th variable with constant C_l on E* if

$$(8.3.1) \quad d_Y(f(x), f(x')) \leq C_l |x_l - x'_l|$$

for every $x, x' \in E$ such that $x_j = x'_j$ when $j \neq l$. This means that $f(x)$ is Lipschitz as a function of x_l with constant C_l , when the other coordinates of x are fixed. More precisely, this uses the restriction of the standard Euclidean metric on \mathbf{R} to the set of real numbers that occur as l th coordinates of elements of E , when the other coordinates are fixed.

This is the same as saying that if L is any line in \mathbf{R}^n that is parallel to the l th coordinate axis, then the restriction of f to $E \cap L$ is Lipschitz with constant C_l .

8.3.1 Functions on cells

Let a_1, \dots, a_n and b_1, \dots, b_n be real numbers such that $a_k < b_k$ for each $k = 1, \dots, n$. Remember that the corresponding set

$$(8.3.2) \quad \mathcal{C} = \{x \in \mathbf{R}^n : a_k \leq x_k \leq b_k \text{ for each } k = 1, \dots, n\}$$

is called a *cell* in \mathbf{R}^n , as in [210]. This is the same as the Cartesian product of the closed intervals $[a_k, b_k]$, $k = 1, \dots, n$, as mentioned in Subsection A.9.1.

Suppose now that f is a mapping from \mathcal{C} into Y . Thus $f(x)$ may be considered as a function of x_l on $[a_l, b_l]$, when the other variables are fixed elements of $[a_k, b_k]$, $k \neq l$.

8.3.2 Using partial Lipschitz conditions

Let C_l be a nonnegative real number for each $l = 1, \dots, n$, and suppose that f is partially Lipschitz in the l th variable with constant C_l on \mathcal{C} for each $l = 1, \dots, n$. Let $x, w \in \mathcal{C}$ be given, and observe that

$$(8.3.3) \quad d_Y(f(x), f(w)) \leq \sum_{l=1}^n C_l |x_l - w_l|.$$

To see this, one can go from x to w in n steps, only changing one coordinate in each step. This also uses the triangle inequality to get that the left side of (8.3.3) is less than or equal to the sum of the distances between the values of f at the appropriate pair of points in each step. The distance in the l th step is less than or equal to the l th term in the sum on the right side of (8.3.3) for each $l = 1, \dots, n$, because of the partial Lipschitz condition for f in the l th variable on \mathcal{C} .

It follows that

$$(8.3.4) \quad d_Y(f(x), f(w)) \leq \left(\max_{1 \leq l \leq n} C_l \right) \|x - w\|_1,$$

where $\|\cdot\|_1$ is defined on \mathbf{R}^n as in Subsection 1.3.2. Similarly,

$$(8.3.5) \quad d_Y(f(x), f(w)) \leq \left(\sum_{l=1}^n C_l^2 \right)^{1/2} \|x - w\|_2,$$

where $\|\cdot\|_2$ is the standard Euclidean norm on \mathbf{R}^n . This uses the Cauchy-Schwarz inequality, applied to the sum on the right side of (8.3.3). We also get that

$$(8.3.6) \quad d_Y(f(x), f(w)) \leq \left(\sum_{l=1}^n C_l \right) \|x - w\|_\infty,$$

where $\|\cdot\|_\infty$ is as defined in Subsection 1.3.2. In particular, f is Lipschitz with respect to the restrictions of the metrics on \mathbf{R}^n associated to these norms to \mathcal{C} .

8.3.3 Functions on some other subsets of \mathbf{R}^n

Let f be a mapping from a set $E \subseteq \mathbf{R}^n$ into Y again, and suppose that f is partially Lipschitz in the l th variable on E with constant $C_l \geq 0$ for every $l = 1, \dots, n$. If \mathcal{C} is a cell in \mathbf{R}^n with

$$(8.3.7) \quad \mathcal{C} \subseteq E,$$

then the restriction of f to \mathcal{C} has the same properties mentioned in the previous subsection. In particular, if $E = \mathbf{R}^n$, then these properties hold for every $x, w \in \mathbf{R}^n$.

8.4 Partial derivatives

Let m and n be positive integers, and let U be a nonempty open subset of \mathbf{R}^n , with respect to the standard Euclidean metric. Also let f be a mapping from U into \mathbf{R}^m , and let $x \in U$ be given.

If l is a positive integer less than or equal to n , then the l th partial derivative

$$(8.4.1) \quad \partial_l f(x) = D_l f(x) = \frac{\partial f}{\partial x_l}(x)$$

of f at x can be defined, as usual, as the derivative of f in the l th variable at x_l , when it exists, and with the other variables being fixed. More precisely, if we consider f as a function of the l th variable, with the k th variable equal to x_k when $k \neq l$, then f is defined on an open subset of the real line that contains x_l .

Note that $\partial_l f(x)$ exists if and only if the l th partial derivative $\partial_l f_j(x)$ of the j th component f_j of f at x exists for each $j = 1, \dots, m$, in which case $\partial_l f_j(x)$ is the same as the j th component of $\partial_l f(x)$ for each $j = 1, \dots, m$, as an element of \mathbf{R}^m .

8.4.1 Partial derivatives on cells

Let a_1, \dots, a_n and b_1, \dots, b_n be real numbers with $a_k < b_k$ for each $k = 1, \dots, n$, and let \mathcal{C} be the corresponding cell in \mathbf{R}^n , as in (8.3.2). Suppose now that f is a mapping from \mathcal{C} into \mathbf{R}^m . Let $x \in \mathcal{C}$ and a positive integer $l \leq n$ be given.

If $a_l < x_l < b_l$, then the l th partial derivative of f at x can be defined as the derivative of f as a function of the l th variable on $[a_l, b_l]$ at x_l , when it exists, and with the other variables being fixed. If $x_l = a_l$ or b_l , then one can use the corresponding one-sided derivative, as before.

Suppose for the moment that f is defined on an open set $U \subseteq \mathbf{R}^n$, and that $\mathcal{C} \subseteq U$. If the l th partial derivative of f at x as a function on U exists, then the l th partial derivative of f at x as a function on \mathcal{C} exists, as in the preceding paragraph, and with the same value.

8.4.2 Partial Lipschitz conditions on cells

Suppose that the l th partial derivative of f exists everywhere on \mathcal{C} , and that it is continuous as a mapping from \mathcal{C} into \mathbf{R}^m , with respect to the restriction of the standard Euclidean metric on \mathbf{R}^n to \mathcal{C} , and the standard Euclidean metric on \mathbf{R}^m .

Let N be a norm on \mathbf{R}^m . Remember that N is continuous on \mathbf{R}^m , as in Subsection 6.11.1, so that $N(\partial_l f(x))$ is continuous as a real-valued function on \mathcal{C} . This implies that $N(\partial_l f(x))$ is bounded on \mathcal{C} , because \mathcal{C} is compact, and that its maximum on \mathcal{C} is attained.

Let d_N be the metric on \mathbf{R}^m associated to N . Using the remarks in Section 8.2, we get that f is partially Lipschitz in the l th variable on \mathcal{C} , with respect to d_N on \mathbf{R}^m , with constant

$$(8.4.2) \quad C_l = \sup_{x \in \mathcal{C}} N(\partial_l f(x)).$$

This is the smallest value of C_l with this property, for essentially the same reasons as before.

8.4.3 Continuous differentiability on open subsets of \mathbf{R}^n

Let f be a mapping from an open subset U of \mathbf{R}^n into \mathbf{R}^m again. Let us say that f is *continuously differentiable* on U if for each $l = 1, \dots, n$, the l th partial derivative $\partial_l f(x)$ of f exists at every $x \in U$, and defines a continuous mapping from U into \mathbf{R}^m . This uses the restriction of the standard Euclidean metric on \mathbf{R}^n to U , and the standard Euclidean metric on \mathbf{R}^m , as usual.

If \mathcal{C} is a cell in \mathbf{R}^n and $\mathcal{C} \subseteq U$, then the restriction of f to \mathcal{C} satisfies the analogous continuous differentiability property on \mathcal{C} .

8.5 A more precise version

Let n be a positive integer, let a_1, \dots, a_n and b_1, \dots, b_n be real numbers with $a_k < b_k$ for each $k = 1, \dots, n$, and let \mathcal{C} be the corresponding cell in \mathbf{R}^n again,

as in (8.3.2). Also let f be a real-valued function on \mathcal{C} , and let $1 \leq l \leq n$ be given. Suppose that $f(x)$ is continuous as a real-valued function of x_l on $[a_l, b_l]$, when $x_k \in [a_k, b_k]$ is kept fixed for $k \neq l$.

Suppose in addition that $f(x)$ is differentiable as a function of x_l on (a_l, b_l) when $x_k \in [a_k, b_k]$ is kept fixed for $k \neq l$, so that $\partial_l f(x)$ is defined under these conditions. Let us suppose as well that $\partial_l f(x)$ is bounded, and put

$$(8.5.1) \quad C_l = \sup\{|\partial_l f(x)| : x \in \mathcal{C}, a_l < x_l < b_l\}.$$

Under these conditions, one can use the mean value theorem to get that f is partially Lipschitz in the l th variable on \mathcal{C} , with respect to the standard Euclidean metric on \mathbf{R} , with constant C_l . This is the smallest value of C_l with this property, as usual.

8.5.1 Real-valued functions on open subsets of \mathbf{R}^n

Now let U be a nonempty open subset of \mathbf{R}^n , with respect to the standard Euclidean metric, and let f be a real-valued function on U . Suppose that $\partial_l f(x)$ exists for every $x \in U$, which implies that f is continuous as a function of x_l , with the other variables kept fixed, everywhere on U . If

$$(8.5.2) \quad \mathcal{C} \subseteq U$$

and $\partial_l f$ is bounded on \mathcal{C} , then we get that f is partially Lipschitz in the l th variable on \mathcal{C} , with constant C_l as in (8.5.1), as before.

If $\partial_l f$ is bounded on U , then this holds for every cell \mathcal{C} as in (8.5.2). If $\partial_l f$ exists and is bounded on U for each $l = 1, \dots, n$, then one can use this to get that f is continuous on U , as in Exercise 7 on p239 of [210].

8.6 Directional derivatives

Let m and n be positive integers, and let U be an open subset of \mathbf{R}^n , with respect to the standard Euclidean metric. Also let f be a mapping from U into \mathbf{R}^m , and let $x \in U$ and $v \in \mathbf{R}^n$ be given.

It is easy to see that

$$(8.6.1) \quad U(x, v) = \{t \in \mathbf{R} : x + tv \in U\}$$

is an open set in the real line, with respect to the standard Euclidean metric on \mathbf{R} . Of course, $0 \in U(x, v)$, because $x \in U$.

Let us consider

$$(8.6.2) \quad f_{x,v}(t) = f(x + tv)$$

as an \mathbf{R}^m -valued function of $t \in U(x, v)$. If this function is differentiable at 0, then put

$$(8.6.3) \quad D_v f(x) = f'_{x,v}(0).$$

This is the *directional derivative* of f at x in the direction v .

8.6.1 A homogeneity property of the directional derivative

Let $r \in \mathbf{R}$ be given, so that $rv \in \mathbf{R}^n$ too. Observe that

$$(8.6.4) \quad U(x, rv) = \{t \in \mathbf{R} : rt \in U(x, v)\},$$

and

$$(8.6.5) \quad f_{x,rv}(t) = f_{x,v}(rt)$$

on $U(x, rv)$. If $D_v f(x)$ exists, then it is easy to see that $D_{rv} f(x)$ exists, with

$$(8.6.6) \quad D_{rv} f(x) = r D_v f(x).$$

8.6.2 Some remarks about directional derivatives

Let e_1, \dots, e_n be the usual standard basis vectors in \mathbf{R}^n , so that the l th coordinate of e_k is equal to 1 when $k = l$, and to 0 when $k \neq l$. The directional derivative

$$(8.6.7) \quad D_{e_k} f(x)$$

is the same as the k th partial derivative $\partial_k f(x)$, when it exists, for each $k = 1, \dots, n$.

If $j \in \{1, \dots, m\}$, then let $f_j(x) \in \mathbf{R}$ be the j th coordinate of $f(x)$, as an element of \mathbf{R}^m . Thus $f_j(x)$ defines a real-valued function on U for each $j = 1, \dots, m$. Of course, the directional derivative $D_v f(x)$ exists if and only if the directional derivative $D_v f_j(x)$ exists for every $j = 1, \dots, m$, in which case $D_v f_j(x)$ is the j th coordinate of $D_v f(x)$ for each j .

Similarly, note that f is continuously differentiable on U if and only if for each $j = 1, \dots, m$, f_j is continuously-differentiable as a real-valued function on U .

8.6.3 Linearity in v

Suppose for the moment that $D_v f(x)$ exists for every $v \in \mathbf{R}^n$. In some situations, we may also have that

$$(8.6.8) \quad D_{v+w} f(x) = D_v f(x) + D_w f(x)$$

for every $v, w \in \mathbf{R}^n$. This means that

$$(8.6.9) \quad v \mapsto D_v f(x)$$

defines a linear mapping from \mathbf{R}^n into \mathbf{R}^m , because of (8.6.6). In this case, we get that

$$(8.6.10) \quad D_v f(x) = \sum_{k=1}^n v_k D_{e_k} f(x) = \sum_{k=1}^n v_k \partial_k f(x)$$

for every $v \in \mathbf{R}^n$, because $v = \sum_{k=1}^n v_k e_k$.

8.7 Differentiable mappings

Let m and n be positive integers again, and let U be a nonempty open subset of \mathbf{R}^n , with respect to the standard Euclidean metric on \mathbf{R}^n . Also let f be a mapping from U into \mathbf{R}^m , and let $x \in U$ be given.

We say that f is *differentiable* at x if there is a linear mapping A from \mathbf{R}^n into \mathbf{R}^m such that

$$(8.7.1) \quad \lim_{h \rightarrow 0} \frac{\|f(x+h) - f(x) - A(h)\|_{2,\mathbf{R}^m}}{\|h\|_{2,\mathbf{R}^n}} = 0.$$

Here $\|\cdot\|_{2,\mathbf{R}^m}$ and $\|\cdot\|_{2,\mathbf{R}^n}$ are the standard Euclidean norms on \mathbf{R}^m and \mathbf{R}^n , respectively. It is easy to see that this implies that f is continuous at x , and some more precise statements along these lines will be discussed in the next section.

8.7.1 Some reformulations of differentiability

Put

$$(8.7.2) \quad U_x = \{h \in \mathbf{R}^n : x+h \in U\},$$

which is an open set in \mathbf{R}^n that contains 0. Of course, $f(x+h)$ is defined for every $h \in U_x$, by construction. Put

$$(8.7.3) \quad a(h) = f(x+h) - f(x) - A(h)$$

for each $h \in U_x$, so that

$$(8.7.4) \quad f(x+h) = f(x) + A(h) + a(h)$$

for every $h \in U_x$. Using this, (8.7.1) says that

$$(8.7.5) \quad \lim_{h \rightarrow 0} \frac{\|a(h)\|_{2,\mathbf{R}^m}}{\|h\|_{2,\mathbf{R}^n}} = 0.$$

Similarly, if $h \in U_x$ and $h \neq 0$, then put

$$(8.7.6) \quad \alpha(h) = a(h) \|h\|_{2,\mathbf{R}^n}^{-1}.$$

Let us put $\alpha(0) = 0$, so that

$$(8.7.7) \quad f(x+h) = f(x) + A(h) + \alpha(h) \|h\|_{2,\mathbf{R}^n}$$

for every $h \in U_x$, as in (8.7.4). Clearly (8.7.5) is equivalent to

$$(8.7.8) \quad \lim_{h \rightarrow 0} \|\alpha(h)\|_{2,\mathbf{R}^m} = 0.$$

8.7.2 The differential of f at the point x

One can check directly that A is unique, when it exists. This can also be obtained from the remarks in the next subsection.

In this case, we put $f'(x) = A$. This may be called the *differential* of f at x .

If $n = 1$, then this reduces to the usual definition of the derivative of a function of one variable. More precisely, a linear mapping from \mathbf{R} into \mathbf{R}^m corresponds to multiplying a real number by a fixed element of \mathbf{R}^m . The differential of f at x is the linear mapping that corresponds to multiplying a real number by the usual derivative of f at x , as an element of \mathbf{R}^m .

8.7.3 Directional derivatives of differentiable mappings

Suppose that f is differentiable at x . If $v \in \mathbf{R}^n$, then one can verify that the directional derivative of f at x in the direction v exists, with

$$(8.7.9) \quad D_v f(x) = f'(x)(v).$$

In particular, if $k \in \{1, \dots, n\}$, then the k th partial derivative of f at x exists, with

$$(8.7.10) \quad \frac{\partial f}{\partial x_k}(x) = f'(x)(e_k).$$

Here e_k is the k th standard basis vector in \mathbf{R}^n , as before.

Remember that linear mappings from \mathbf{R}^n into \mathbf{R}^m are associated to $m \times n$ matrices of real numbers in a standard way, using the standard basis vectors in \mathbf{R}^n . The entries of the matrix associated to $f'(x)$ are given by the partial derivatives of the m components of f at x . If $m = 1$, then this is related to the *gradient* of f at x , as in (8.6.10).

8.7.4 Differentiability of sums and products

If $t \in \mathbf{R}$, then tf is a function defined on U with values in \mathbf{R}^m too. It is easy to see tf is differentiable at x as well, with

$$(8.7.11) \quad (tf)'(x) = t f'(x).$$

Similarly, let g be another mapping from U into \mathbf{R}^m that is differentiable at x . One can check that $f + g$ is differentiable at x , with

$$(8.7.12) \quad (f + g)'(x) = f'(x) + g'(x).$$

One can verify that a mapping from U into \mathbf{R}^m is differentiable at x if and only if its m components are differentiable at x as real-valued functions on U .

Suppose now that f, g are real-valued functions on U that are differentiable at x . It is not too difficult to show that fg is differentiable at x as well, with

$$(8.7.13) \quad (fg)'(x) = g(x) f'(x) + f(x) g'(x).$$

This can also be obtained from the chain rule, as in Subsection 8.9.6.

8.8 Pointwise Lipschitz conditions

Let (X, d_X) and (Y, d_Y) be metric spaces, and let f be a mapping from X into Y . Also let x be an element of X , let C be a nonnegative real number, and let r be a positive real number.

Let us say that

(8.8.1) f is *pointwise Lipschitz at x with constant C up to the scale r*

if for every $w \in X$ with

$$(8.8.2) \quad d_X(x, w) < r,$$

we have that

$$(8.8.3) \quad d_Y(f(x), f(w)) \leq C d_X(x, w).$$

Of course, this implies that

$$(8.8.4) \quad f \text{ is continuous at } x.$$

We may allow $r = +\infty$ here too, so that (8.8.2) holds for every $w \in X$.

8.8.1 Defining differentiability using other norms

Let m and n be positive integers, let U be a nonempty open subset of \mathbf{R}^n , and let f be a mapping from U into \mathbf{R}^m . Also let $x \in U$ be given, and suppose that

$$(8.8.5) \quad f \text{ is differentiable at } x.$$

If N and N_0 are norms on \mathbf{R}^m and \mathbf{R}^n , respectively, then (8.7.1) is equivalent to

$$(8.8.6) \quad \lim_{h \rightarrow 0} \frac{N(f(x+h) - f(x) - f'(x)(h))}{N_0(h)} = 0.$$

This follows from the comparison between N , N_0 and the standard Euclidean norms on \mathbf{R}^m , \mathbf{R}^n , respectively, as in Subsections 6.10.3 and 6.11.3.

Let $\epsilon > 0$ be given, so that there is a $\delta > 0$ such that

$$(8.8.7) \quad \frac{N(f(x+h) - f(x) - f'(x)(h))}{N_0(h)} < \epsilon$$

for every $h \in \mathbf{R}^n$ such that $h \neq 0$,

$$(8.8.8) \quad N_0(h) < \delta,$$

and $x+h \in U$. This uses the comparison between N_0 and the standard Euclidean norm on \mathbf{R}^n again, to express (8.8.8) in terms of $N_0(h)$. We may as well take δ small enough so that (8.8.8) implies that $x+h \in U$.

8.8.2 Some pointwise Lipschitz conditions

If $h \in \mathbf{R}^n$ satisfies (8.8.8), then it follows that

$$(8.8.9) \quad N(f(x+h) - f(x) - f'(x)(h)) \leq \epsilon N_0(h).$$

This implies that

$$(8.8.10) \quad \begin{aligned} N(f(x+h) - f(x)) &\leq N(f(x+h) - f(x) - f'(x)(h)) + N(f'(x)(h)) \\ &\leq N(f'(x)(h)) + \epsilon N_0(h) \end{aligned}$$

when (8.8.8) holds, using the triangle inequality for N in the first step. Using this, we get that

$$(8.8.11) \quad \begin{aligned} N(f(x+h) - f(x)) &\leq \|f'(x)\|_{op} N_0(h) + \epsilon N_0(h) \\ &= (\|f'(x)\|_{op} + \epsilon) N_0(h) \end{aligned}$$

when (8.8.8) holds, where $\|f'(x)\|_{op}$ is the operator norm of $f'(x)$ corresponding to N_0 and N , as in Section 7.5. This is a pointwise Lipschitz condition for f at x , with respect to the metrics associated to N_0 and N .

8.8.3 Estimating $\|f'(x)\|_{op}$

Using (8.8.9) and the triangle inequality for N again, we get that

$$(8.8.12) \quad N(f'(x)(h)) \leq N(f(x+h) - f(x)) + \epsilon N_0(h)$$

when $h \in \mathbf{R}^n$ satisfies (8.8.8).

Let r be a positive real number, let $C(r)$ be a nonnegative real number, and suppose that

$$(8.8.13) \quad f \text{ is pointwise Lipschitz at } x \text{ with constant } C(r) \text{ up to scale } r,$$

with respect to the metric on \mathbf{R}^m associated to N , and the restriction to U of the metric on \mathbf{R}^n associated to N_0 . This means that

$$(8.8.14) \quad N(f(x+h) - f(x)) \leq C(r) N_0(h)$$

for every $h \in \mathbf{R}^n$ such that

$$(8.8.15) \quad N_0(h) < r$$

and $x+h \in U$. As before, we may as well take r small enough so that (8.8.15) implies that $x+h \in U$, using the comparison between N_0 and the standard Euclidean norm on \mathbf{R}^n .

Combining (8.8.12) and (8.8.14), we get that

$$(8.8.16) \quad N(f'(x)(h)) \leq C(r) N_0(h) + \epsilon N_0(h) = (C(r) + \epsilon) N_0(h)$$

for every $h \in \mathbf{R}^n$ with

$$(8.8.17) \quad N_0(h) < \min(\delta, r).$$

It is easy to see that this implies that (8.8.16) holds for every $h \in \mathbf{R}^n$, because $f'(x)(h)$ is linear in h . More precisely, one can reduce to the case where (8.8.17) holds, by multiplying h by a sufficiently small positive real number.

It follows that

$$(8.8.18) \quad \|f'(x)\|_{op} \leq C(r) + \epsilon.$$

This implies that

$$(8.8.19) \quad \|f'(x)\|_{op} \leq C(r),$$

because $\epsilon > 0$ is arbitrary.

8.9 The chain rule

Let m , n , and p be positive integers, let U be a nonempty open subset of \mathbf{R}^n , and let f be a mapping from U into \mathbf{R}^m . Also let V be an open set in \mathbf{R}^m , and suppose that

$$(8.9.1) \quad f(U) \subseteq V.$$

If g is a mapping from V into \mathbf{R}^p , then the composition $g \circ f$ of f and g is defined as a mapping from U into \mathbf{R}^p , i.e., $(g \circ f)(x) = g(f(x))$ for all $x \in U$.

Suppose that

$$(8.9.2) \quad f \text{ is differentiable at } x \in U,$$

and that

$$(8.9.3) \quad g \text{ is differentiable at } f(x) \in V.$$

Thus the differential $f'(x)$ of f at x is defined as a linear mapping from \mathbf{R}^n into \mathbf{R}^m , and the differential $g'(f(x))$ of g at $f(x)$ is defined as a linear mapping from \mathbf{R}^m into \mathbf{R}^p .

8.9.1 The precise statement of the chain rule

Under these conditions, it is not too difficult to show that

$$(8.9.4) \quad g \circ f \text{ is differentiable at } x,$$

with

$$(8.9.5) \quad (g \circ f)'(x) = g'(f(x)) \circ f'(x).$$

This is the analogue of the chain rule in this situation.

One can check that this reduces to the usual version of the chain rule for real-valued functions of one real variable when $m = n = p = 1$.

8.9.2 A way to get the chain rule

If g is a linear mapping from \mathbf{R}^m into \mathbf{R}^p , then one can check the chain rule directly. Otherwise, one can use this to reduce to the case where

$$(8.9.6) \quad g'(f(x)) = 0.$$

In this case, one can verify that

$$(8.9.7) \quad (g \circ f)'(x) = 0$$

when f satisfies a pointwise Lipschitz condition at x .

8.9.3 Partial derivatives of compositions

In particular, (8.9.5) can be used to obtain the partial derivatives of $g \circ f$ at x in terms of the partial derivatives of f at x and the partial derivatives of g at $f(x)$. Similarly, let $v \in \mathbf{R}^n$ be given, so that

$$(8.9.8) \quad w = f'(x)(v) = D_v f(x)$$

is an element of \mathbf{R}^m . Using (8.9.5), we get that

$$(8.9.9) \quad \begin{aligned} D_v(g \circ f)(x) &= (g \circ f)'(x)(v) &= g'(f(x))(f'(x)(v)) \\ &= g'(f(x))(w) = (D_w g)(f(x)). \end{aligned}$$

If $m = 1$, then this can be obtained from the usual chain rule for differentiable real-valued functions of one real variable.

8.9.4 Constant functions and linear mappings

Let l_1, l_2 be positive integers. Of course, constant mappings from \mathbf{R}^{l_1} into \mathbf{R}^{l_2} are differentiable at every point in \mathbf{R}^{l_1} , with differential equal to 0.

If $f = A$ is a linear mapping from \mathbf{R}^{l_1} into \mathbf{R}^{l_2} , then f is differentiable at every point $x \in \mathbf{R}^{l_1}$, with $f'(x) = A$.

8.9.5 Squares of differentiable functions

It is well known that $t \mapsto t^2$ is differentiable as a real-valued function on the real line, with derivative equal to $2t$. Let f be a real-valued function on U , and suppose that f is differentiable at $x \in U$. Using the chain rule, we get that

$$(8.9.10) \quad f^2 \text{ is differentiable at } x$$

too, with

$$(8.9.11) \quad (f^2)'(x) = 2f(x)f'(x).$$

8.9.6 Products of differentiable functions

Let g be another real-valued function on U that is differentiable at x , so that g^2 is differentiable at x as well, with

$$(8.9.12) \quad (g^2)'(x) = 2g(x)g'(x).$$

Similarly, $f + g$ is differentiable at x , so that $(f + g)^2$ is differentiable at x , with

$$(8.9.13) \quad ((f + g)^2)'(x) = 2(f(x) + g(x))(f'(x) + g'(x)).$$

Note that

$$(8.9.14) \quad fg = (1/2)((f + g)^2 - f^2 - g^2)$$

on U . It follows that fg is differentiable at x , with $(fg)'(x)$ as in (8.7.13).

8.9.7 Reciprocals of differentiable functions

Similarly, it is well known that $t \mapsto 1/t$ is differentiable as a real-valued function on $\mathbf{R} \setminus \{0\}$, with derivative equal to $-1/t^2$. Let f be a real-valued function on U such that $f(w) \neq 0$ for every $w \in U$, and suppose that f is differentiable at $x \in U$. The chain rule implies that

$$(8.9.15) \quad 1/f \text{ is differentiable at } x,$$

with

$$(8.9.16) \quad (1/f)'(x) = (-1/f(x)^2)f'(x).$$

8.9.8 Polynomials and rational functions

In particular, one can use these remarks to get that polynomial functions on \mathbf{R}^n are differentiable at every point. Similarly, rational functions are differentiable on open sets where the denominator is not zero.

8.10 Continuous differentiability

Let m and n be positive integers, let U be an open set in \mathbf{R}^n , and let f be a mapping from U into \mathbf{R}^m . If

$$(8.10.1) \quad f \text{ is differentiable at every point in } U,$$

then f is said to be *differentiable* on U .

Remember that f is said to be continuously differentiable on U if for each $k = 1, \dots, n$, the k th partial derivative $\partial_k f(x)$ exists at every $x \in U$, and is continuous as a mapping from U into \mathbf{R}^m , as in Subsection 8.4.3. It is well known that

$$(8.10.2) \quad \text{continuously-differentiable mappings on } U \text{ are differentiable}$$

in the sense just mentioned. More precisely, if $x \in U$, then $f'(x)$ is the linear mapping from \mathbf{R}^n into \mathbf{R}^m defined by

$$(8.10.3) \quad f'(x)(h) = \sum_{k=1}^n h_k \partial_k f(x)$$

for every $h \in \mathbf{R}^n$.

8.10.1 The case where $\partial_k f(x) = 0$, $1 \leq k \leq n$

To see this, suppose that f is continuously differentiable on U , and let $x \in U$ be given. Suppose for the moment that

$$(8.10.4) \quad \partial_k f(x) = 0 \quad \text{for each } k = 1, \dots, n,$$

so that we would like to show that f is differentiable at x , with $f'(x) = 0$. This is equivalent to saying that

$$(8.10.5) \quad \lim_{h \rightarrow 0} \frac{\|f(x+h) - f(x)\|_{2, \mathbf{R}^m}}{\|h\|_{2, \mathbf{R}^n}} = 0.$$

Because $\partial_k f$ is continuous on U , (8.10.4) implies that $\partial_k f$ is as small as we want near x . One can use this to get (8.10.5), as in Subsections 8.3.2 and 8.4.2.

Alternatively, it is not difficult to reduce to the case where $m = 1$. This permits one to use the mean value theorem, as in Section 8.5. In order to estimate

$$(8.10.6) \quad f(x+h) - f(x),$$

one can go from x to $x+h$ in n steps, only changing one coordinate in each step, as in Subsection 8.3.2.

8.10.2 Reducing to the previous case

We can reduce to the case where (8.10.4) holds, as follows. If $w \in U$, then put

$$(8.10.7) \quad \tilde{f}(w) = f(w) - \sum_{k=1}^n (w_k - x_k) \partial_k f(x).$$

This defines a mapping from U into \mathbf{R}^m , as a function of w . The partial derivatives of \tilde{f} are given by

$$(8.10.8) \quad \partial_l \tilde{f}(w) = \partial_l f(w) - \partial_l f(x)$$

for every $w \in U$ and $l = 1, \dots, n$. In particular, \tilde{f} is continuously differentiable on U , with

$$(8.10.9) \quad \partial_l \tilde{f}(x) = \partial_l f(x) - \partial_l f(x) = 0$$

for every $l = 1, \dots, n$.

This implies that

$$(8.10.10) \quad \lim_{h \rightarrow 0} \frac{\|\tilde{f}(x+h) - \tilde{f}(x)\|_{2, \mathbf{R}^n}}{\|h\|_{2, \mathbf{R}^m}} = 0,$$

as before. This is the same as saying that f is differentiable at x , with $f'(x)$ as in (8.10.3).

8.10.3 Another formulation of continuous differentiability

We can think of $f'(x)$ as a function of $x \in U$ with values in the space $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ of linear mappings from \mathbf{R}^n into \mathbf{R}^m . One can get nice metrics on $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ from norms in the usual way, such as the Hilbert–Schmidt norm or the operator norm associated to the standard Euclidean norms on \mathbf{R}^n and \mathbf{R}^m .

If f is continuously differentiable on U , then one can check that

$$(8.10.11) \quad f' \text{ is continuous as a mapping from } U \text{ into } \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m),$$

with respect to such a metric. Conversely, if f is differentiable on U , and if f' is continuous on U in this sense, then one can verify that

$$(8.10.12) \quad \text{the partial derivatives of } f \text{ are continuous on } U.$$

Thus continuous differentiability on U can be defined equivalently in this way.

8.11 Another basic Lipschitz estimate

Let m and n be positive integers, let U be a nonempty open subset of \mathbf{R}^n , and let f be a continuously-differentiable mapping from U into \mathbf{R}^m . Thus f is differentiable at every $x \in U$, as in the previous section, so that $f'(x)$ is defined as a linear mapping from \mathbf{R}^n into \mathbf{R}^m for every $x \in U$. We also have that f' is continuous on U , as before.

8.11.1 Continuity of $\|f'(x)\|_{op}$

Let N and N_0 be norms on \mathbf{R}^m and \mathbf{R}^n , respectively, and let $\|\cdot\|_{op} = \|\cdot\|_{op, N_0 N}$ be the corresponding operator norm for linear mappings from \mathbf{R}^n into \mathbf{R}^m , as in Section 7.5. Remember that N and N_0 can be compared with the standard Euclidean norms on \mathbf{R}^m and \mathbf{R}^n , respectively, as in Subsections 6.10.3 and 6.11.3. This leads to an analogous comparison between $\|\cdot\|_{op}$ and the operator norm for linear mappings from \mathbf{R}^n into \mathbf{R}^m with respect to the standard Euclidean norms on \mathbf{R}^n and \mathbf{R}^m .

In particular, one can use this to check that $\|\cdot\|_{op}$ is continuous on the space of linear mappings from \mathbf{R}^n into \mathbf{R}^m , with respect to the metric that corresponds to the Hilbert–Schmidt norm on $m \times n$ matrices of real numbers. This implies that

$$(8.11.1) \quad \|f'(x)\|_{op} \text{ is continuous as a real-valued function on } U,$$

because $f'(x)$ is continuous on U .

If $x \in U$ and $v \in \mathbf{R}^n$, then the directional derivative $D_v f(x)$ of f at x in the direction of v is the same as $f'(x)$ applied to v , as in (8.7.9). Hence

$$(8.11.2) \quad N(D_v f(x)) = N(f'(x)(v)) \leq \|f'(x)\|_{op} N_0(v),$$

by the definition of $\|\cdot\|_{op}$.

8.11.2 Convex open sets U

Let us suppose from now on in this section that U is also convex in \mathbf{R}^n . This means that for each $x, w \in U$ and $t \in \mathbf{R}$ with $0 \leq t \leq 1$ we have that

$$(8.11.3) \quad tx + (1-t)w \in U.$$

Let $x, w \in U$ be given, and consider

$$(8.11.4) \quad f(tx + (1-t)w) = f(w + t(x-w))$$

as a function of $t \in [0, 1]$ with values in \mathbf{R}^m . Because f is differentiable on U , we can differentiate (8.11.4) in t , to get that

$$(8.11.5) \quad \frac{d}{dt}(f(tx + (1-t)w)) = (D_{(x-w)}f)(tx + (1-t)w)$$

for every $t \in [0, 1]$. It follows that

$$(8.11.6) \quad \int_0^1 (D_{(x-w)}f)(tx + (1-t)w) dt = f(x) - f(w),$$

by the fundamental theorem of calculus.

8.11.3 Estimating $N(f(x) - f(w))$

This implies that

$$(8.11.7) \quad N(f(x) - f(w)) \leq \int_0^1 N((D_{(x-w)}f)(tx + (1-t)w)) dt,$$

as in Subsection 8.1.3. Combining this with (8.11.2), we get that

$$(8.11.8) \quad N(f(x) - f(w)) \leq N_0(x-w) \int_0^1 \|f'(tx + (1-t)w)\|_{op} dt.$$

Suppose that f' is bounded on U , and observe that

$$(8.11.9) \quad N(f(x) - f(w)) \leq \left(\sup_{u \in U} \|f'(u)\|_{op} \right) N_0(x-w),$$

by (8.11.8). This means that f is Lipschitz as a mapping from U into \mathbf{R}^m , using the restriction to U of the metric on \mathbf{R}^n associated to N_0 , and the metric on \mathbf{R}^m associated to N .

More precisely, the corresponding Lipschitz constant is given by

$$(8.11.10) \quad \text{Lip}(f) = \sup_{u \in U} \|f'(u)\|_{op}.$$

Indeed, $\text{Lip}(f)$ is less than or equal to the right side of (8.11.10), because of (8.11.9). In order to get the opposite inequality, one can check directly that $\|f'(u)\|_{op}$ is less than or equal to $\text{Lip}(f)$ for every $u \in U$, as in Subsection 8.8.3.

8.11.4 The $m = 1$ case

If $m = 1$, then we can take N to be the usual absolute value function on \mathbf{R} . In this case, one can use the mean value theorem to get that

$$(8.11.11) \quad |f(x) - f(w)| \leq \left(\sup_{u \in U} \|f'(u)\|_{op} \right) N_0(x - w)$$

for every $x, w \in U$ when f is differentiable on U , and f' is bounded on U . More precisely, one can consider (8.11.4) as a real-valued function of $t \in [0, 1]$, whose derivative can be estimated in the same way as before. The corresponding Lipschitz constant of f on U is given by (8.11.10), as before.

There are some analogous remarks when $m \geq 2$, as in Subsection 8.2.4.

8.12 Some remarks about connectedness

Let (X, d_X) be a metric space. Remember that subsets A, B of X are said to be *separated* in X if

$$(8.12.1) \quad \overline{A} \cap B = A \cap \overline{B} = \emptyset,$$

where $\overline{A}, \overline{B}$ are the closures of A, B in X , respectively. A subset E of X is said to be *connected* if E cannot be expressed as the union of two nonempty separated subsets of X , as usual.

8.12.1 Subsets of $X_0 \subseteq X$

Let X_0 be a subset of X , and remember that X_0 may be considered as a metric space with respect to the restriction of $d_X(x, w)$ to $x, w \in X_0$. If $A \subseteq X_0$, then let \overline{A}_X be the closure of A in X , and let \overline{A}_{X_0} be the closure of A in X_0 .

One can check that

$$(8.12.2) \quad \overline{A}_{X_0} = \overline{A}_X \cap X_0$$

for every $A \subseteq X_0$. More precisely, if x is any element of X_0 , then one can verify that x is a limit point of A as a subset of X if and only if x is a limit point of A as a subset of X_0 .

If A, B are subsets of X_0 , then one can use this to verify that

$$(8.12.3) \quad A, B \text{ are separated as subsets of } X_0$$

if and only if

$$(8.12.4) \quad A, B \text{ are separated as subsets of } X.$$

If E is a subset of X_0 , then the previous statement implies that

$$(8.12.5) \quad \begin{aligned} E \text{ is connected as a subset of } X_0 \text{ if and only if} \\ E \text{ is connected as a subset of } X. \end{aligned}$$

8.12.2 Disjoint open and closed sets

Note that disjoint closed subsets of X are separated in X . One can check that

$$(8.12.6) \quad \text{disjoint open subsets of } X \text{ are separated in } X$$

too, because the complement of an open set is a closed set.

If X is not connected, then X can be expressed as the union of two nonempty separated sets A, B . In this case, one can verify that

$$(8.12.7) \quad A \text{ and } B \text{ are closed sets in } X.$$

Equivalently, this means that

$$(8.12.8) \quad A \text{ and } B \text{ are open sets in } X$$

under these conditions.

8.12.3 Connected open and closed sets

Suppose that $E \subseteq X$ is not connected, so that $E = A \cup B$ for some nonempty separated subsets A, B of X . If E is a closed set in X , then one can check that (8.12.7) holds. This can be obtained directly from the definitions, or by reducing to the case where $X = E$.

Similarly, if E is an open set in X , then one can verify that (8.12.8) holds.

8.13 Locally constant mappings

Let (X, d_X) be a metric space, let Y be a set, and let f be a mapping from X into Y . We say that f is *locally constant* on X if for every $x \in X$ there is a positive real number $r(x)$ such that

$$(8.13.1) \quad f(w) = f(x)$$

for every $w \in X$ with $d_X(x, w) < r(x)$.

If Y is equipped with a metric d_Y , then

(8.13.2) any locally constant mapping from X into Y is continuous.

If Y is equipped with the discrete metric, as mentioned in Section A.11, then it is easy to see that

(8.13.3) every continuous mapping from X into Y is locally constant.

This also works as long as every element of Y is an isolated point in Y .

8.13.1 Locally constant mappings and connectedness

Let f be a locally constant mapping from X into Y , and let E be a subset of X . If E is connected as a subset of X , then

(8.13.4) f is constant on E .

Equivalently, if f is not constant on E , then E is not connected.

One way to see this is to use the fact that f is continuous with respect to the discrete metric on Y . It is well known that any continuous mapping from X into another metric space maps connected subsets of X to connected subsets of the range. If E is connected in X , then it follows that

(8.13.5) $f(E)$ is connected with respect to the discrete metric on Y .

However, one can check that a subset of Y is connected with respect to the discrete metric if and only if it has at most one element. If $f(E)$ has at most one element, then (8.13.4) holds.

If X is not connected, then X can be expressed as the union of two nonempty disjoint open sets A and B , as in Subsection 8.12.2. One can use this to get a mapping f from X into any set Y with at least two elements such that f is locally constant on X , and not constant on X . One can take f to be constant on each of A and B , with different constant values on these two sets.

8.13.2 The condition $f'(x) \equiv 0$

Let m and n be positive integers, let U be a nonempty open subset of \mathbf{R}^n , and let f be a mapping from U into \mathbf{R}^m . Suppose that f is differentiable on U , and that

(8.13.6) $f'(x) = 0$

for every $x \in U$.

If U is convex, then it follows that f is constant on U , as in Subsection 8.11.2. Similarly,

(8.13.7) the restriction of f to any convex open subset of U is constant.

One can check that open balls in \mathbf{R}^n with respect to the metric associated to any norm on \mathbf{R}^n are convex, directly from the definitions.

This implies that
 (8.13.8) f is locally constant on U ,

with respect to the restriction to U of the standard Euclidean metric on \mathbf{R}^n .

Of course, if f is any mapping from U into \mathbf{R}^m that is locally constant with respect to the restriction to U of the standard Euclidean metric on \mathbf{R}^n , then f is differentiable on U , and satisfies (8.13.6) for every $x \in U$.

If U is also connected as a subset of \mathbf{R}^n , then (8.13.8) implies that

(8.13.9) f is constant on U ,

as in the previous subsection. More precisely, this uses the fact that U is connected as a subset of itself, with respect to the restriction of the standard Euclidean metric on \mathbf{R}^n to U , as in Subsection 8.12.1. This is related to Exercise 9 on p239 of [210].

If U is not connected as a subset of \mathbf{R}^n , then U can be expressed as the union of two nonempty disjoint open subsets of \mathbf{R}^n , as in Subsection 8.12.3. One can use this to get a locally constant mapping from U into \mathbf{R}^m that is not constant on U , as in the previous subsection again.

8.14 Some local Lipschitz conditions

Let (X, d_X) and (Y, d_Y) be metric spaces, and let f be a mapping from X into Y . If $x \in X$ and $r > 0$, then we let

(8.14.1) $B_X(x, r)$

be the open ball in X centered at x with radius r with respect to d_X , as usual. Let us say that

(8.14.2) f is *locally Lipschitz with constant $C \geq 0$ at x*

if there is a positive real number r such that

(8.14.3) the restriction of f to $B_X(x, r)$ is Lipschitz with constant C ,

with respect to the restriction of $d_X(\cdot, \cdot)$ to $B_X(x, r)$. Note that this implies that f is pointwise Lipschitz at x with constant C up to the scale r , as in Section 8.8.

We may also simply say that

(8.14.4) f is *locally Lipschitz at x*

if there is a nonnegative real number C such that f is locally Lipschitz at x with constant C . Similarly, we may say that

(8.14.5) f is *locally Lipschitz on X*

if f is locally Lipschitz at every $x \in X$. Of course, this implies that

$$(8.14.6) \quad f \text{ is continuous on } X.$$

Note that f is locally constant on X exactly when

$$(8.14.7) \quad f \text{ is locally Lipschitz with constant } C = 0 \text{ at every } x \in X.$$

Let W be an open subset of X , and suppose that the restriction of f to W is Lipschitz with constant $C \geq 0$, with respect to the restriction of $d_X(\cdot, \cdot)$ to W . This implies that for each $w \in W$, f is locally Lipschitz at w with constant C .

8.14.1 Local Lipschitz conditions and compactness

Let K be a compact subset of X , and suppose that

$$(8.14.8) \quad \text{for each } x \in K, f \text{ is locally Lipschitz at } x.$$

This means that for every $x \in K$, there is a positive real number $r(x)$ and a nonnegative real number $C(x)$ such that the restriction of f to $B_X(x, r(x))$ is Lipschitz with constant $C(x)$.

Thus K can be covered by open balls of this type, and so there are finitely many elements x_1, \dots, x_l of K such that

$$(8.14.9) \quad K \subseteq \bigcup_{j=1}^l B_X(x_j, r(x_j)),$$

because K is compact. If we put

$$(8.14.10) \quad C = \max_{1 \leq j \leq l} C(x_j),$$

then it is easy to see that

$$(8.14.11) \quad f \text{ is locally Lipschitz with constant } C$$

at every $x \in K$. More precisely, this holds at every x in the right side of (8.14.9).

8.14.2 Some Lipschitz conditions on convex subsets of \mathbf{R}^n

Let m and n be positive integers, let U be a nonempty open subset of \mathbf{R}^n , and let f be a continuously-differentiable mapping from U into \mathbf{R}^m . Also let N and N_0 be norms on \mathbf{R}^m and \mathbf{R}^n , respectively, and let $\|\cdot\|_{op}$ be the corresponding operator norm for linear mappings from \mathbf{R}^n into \mathbf{R}^m .

Suppose that E is a nonempty convex subset of \mathbf{R}^n such that $E \subseteq U$, and that

$$(8.14.12) \quad \|f'(u)\|_{op} \text{ is bounded on } E.$$

If $x, w \in E$, then

$$(8.14.13) \quad N(f(x) - f(w)) \leq \left(\sup_{u \in E} \|f'(u)\|_{op} \right) N_0(x - w),$$

as in Subsection 8.11.3. This means that

$$(8.14.14) \quad \text{the restriction of } f \text{ to } E \text{ is Lipschitz}$$

with respect to the metric on \mathbf{R}^m associated to N , and the restriction to E of the metric on \mathbf{R}^n associated to N_0 .

If E is also closed and bounded in \mathbf{R}^n , then E is compact in \mathbf{R}^n . It is well known that this implies that E is compact as a subset of U , with respect to the restriction to U of the standard Euclidean metric on \mathbf{R}^n . Remember that $\|f'(u)\|_{op}$ is continuous as a real-valued function on U , as in Subsection 8.11.1. It follows that (8.14.12) holds under these conditions.

8.14.3 Some more Lipschitz conditions on subsets of U

One can check that open and closed balls in \mathbf{R}^n with respect to the metric associated to a norm are convex subsets of \mathbf{R}^n , as before. If $v \in U$ and ϵ is a positive real number, then

$$(8.14.15) \quad \|f'(u)\|_{op} < \|f'(v)\|_{op} + \epsilon$$

when $u \in U$ is sufficiently close to v , because $\|f'(u)\|_{op}$ is continuous at v . One can use this to get that f is locally Lipschitz at v with constant

$$(8.14.16) \quad C = \|f'(v)\|_{op} + \epsilon,$$

with respect to the metric on \mathbf{R}^m associated to N , and the restriction to U of the metric on \mathbf{R}^n associated to N_0 .

We also have that $\|f'(u)\|_{op}$ is bounded on any closed ball contained in U , as in the previous subsection. This implies that the restriction of f to such a ball is Lipschitz, as before.

Chapter 9

More on differentiable mappings

9.1 Continuously-differentiable mappings

Let m and n be positive integers, let U be an open subset of \mathbf{R}^n , and let f be a continuously-differentiable mapping from U into \mathbf{R}^m . Also let x be an element of U , and put

$$(9.1.1) \quad A = f'(x),$$

for convenience. Observe that $f - A$ is continuously differentiable as a mapping from U into \mathbf{R}^m too. If $w \in U$, then

$$(9.1.2) \quad (f - A)'(w) = f'(w) - A = f'(w) - f'(x).$$

This tends to 0 as w approaches x , because f is continuously differentiable on U .

9.1.1 Using the operator norm

Let N_0 and N be norms on \mathbf{R}^n and \mathbf{R}^m , respectively. Thus

$$(9.1.3) \quad d_{N_0}(v, w) = N_0(v - w) \text{ and } d_N(y, z) = N(y - z)$$

are metrics on \mathbf{R}^n and \mathbf{R}^m , respectively. The corresponding operator norm for linear mappings from \mathbf{R}^n into \mathbf{R}^m is denoted $\|\cdot\|_{op} = \|\cdot\|_{op, N_0 N}$, as usual. In particular,

$$(9.1.4) \quad \|f'(w) - A\|_{op} = \|f'(w) - f'(x)\|_{op} \rightarrow 0 \quad \text{as } w \rightarrow x,$$

because f is continuously-differentiable on U .

9.1.2 Convex open subsets of U

Let U_1 be a convex open subset of \mathbf{R}^n such that $x \in U_1$ and $U_1 \subseteq U$. Suppose that

$$(9.1.5) \quad f' \text{ is bounded on } U_1,$$

so that

$$(9.1.6) \quad f' - A \text{ is bounded on } U_1$$

as well. Put

$$(9.1.7) \quad b_1 = \sup_{w \in U_1} \|f'(w) - A\|_{op}.$$

Note that

$$(9.1.8) \quad \text{the restriction of } f - A \text{ to } U_1 \text{ is Lipschitz with constant } b_1,$$

as in Subsection 8.11.3 and 8.14.2. This uses the restriction of d_{N_0} to U_1 , and d_N on \mathbf{R}^m .

9.1.3 A lower bound for $d_N(f(u), f(v))$

Suppose that

$$(9.1.9) \quad N(A(v)) = N(f'(x)(v)) \geq c N_0(v)$$

for some $c > 0$ and all $v \in \mathbf{R}^n$. This implies that

$$(9.1.10) \quad d_N(A(u), A(v)) \geq c d_{N_0}(u, v)$$

for every $u, v \in \mathbf{R}^n$, as in (7.9.6). Remember that there is a positive real number c such that (9.1.9) holds for all $v \in \mathbf{R}^n$ exactly when A is one-to-one on \mathbf{R}^n , as in Subsections 7.9.3 and 7.10.1.

Suppose that

$$(9.1.11) \quad b_1 < c.$$

Under these conditions, we have that

$$(9.1.12) \quad d_N(f(u), f(v)) \geq (c - b_1) d_{N_0}(u, v)$$

for every $u, v \in U_1$. More precisely, this corresponds to (7.11.5), with some changes in notation.

9.1.4 A lower bound for $N(f'(w)(v))$

Similarly, if $w \in U_1$, then

$$(9.1.13) \quad N(f'(w)(v)) \geq (c - b_1) N_0(v)$$

for every $v \in \mathbf{R}^n$. This corresponds to (7.11.10).

Suppose for instance that U_1 is the open ball in \mathbf{R}^n centered at x with radius $r > 0$ with respect to d_{N_0} . This is a convex open subset of \mathbf{R}^n , as before. If r is small enough, then $U_1 \subseteq U$, because U is an open set in \mathbf{R}^n that contains x . We can make b_1 as small as we want by taking r small enough, because of (9.1.4). In particular, (9.1.11) holds when r is sufficiently small.

Some more properties of mappings like these will be discussed in Section 9.8.

9.2 Invertible linear mappings

Let n be a positive integer, and let

$$(9.2.1) \quad \mathcal{L}(\mathbf{R}^n) = \mathcal{L}(\mathbf{R}^n, \mathbf{R}^n)$$

be the space of linear mappings from \mathbf{R}^n into itself. Also let N be a norm on \mathbf{R}^n , and let $\|A\|_{op} = \|A\|_{op, NN}$ be the corresponding operator norm for $A \in \mathcal{L}(\mathbf{R}^n)$. More precisely, this uses N on \mathbf{R}^n as both the domain and range of A . Remember that

$$(9.2.2) \quad d_{op}(A, B) = \|A - B\|_{op}$$

defines a metric on $\mathcal{L}(\mathbf{R}^n)$.

Let I be the identity mapping on \mathbf{R}^n , which sends every element of \mathbf{R}^n to itself. It is easy to see that

$$(9.2.3) \quad \|I\|_{op} = 1,$$

as mentioned in Subsection 7.5.4. If $A, B \in \mathcal{L}(\mathbf{R}^n)$, then their composition $B \circ A$ defines a linear mapping from \mathbf{R}^n into itself as well, so that $B \circ A \in \mathcal{L}(\mathbf{R}^n)$. Remember that

$$(9.2.4) \quad \|B \circ A\|_{op} \leq \|A\|_{op} \|B\|_{op},$$

as in Subsection 7.8.3.

A linear mapping A from \mathbf{R}^n into itself is said to be *invertible* if A is a one-to-one mapping from \mathbf{R}^n onto itself. In this case, the inverse mapping A^{-1} is linear as well. The collection of invertible linear mappings on \mathbf{R}^n may be denoted $GL(\mathbf{R}^n)$. It is well known that $A \in \mathcal{L}(\mathbf{R}^n)$ is one-to-one if and only if $A(\mathbf{R}^n) = \mathbf{R}^n$.

9.2.1 The operator norm of A^{-1}

If $A \in GL(\mathbf{R}^n)$, then

$$(9.2.5) \quad N(A^{-1}(u)) \leq \|A^{-1}\|_{op} N(u)$$

for every $u \in \mathbf{R}^n$. More precisely, $\|A^{-1}\|_{op}$ is the smallest nonnegative real number with this property, by definition of the operator norm.

Equivalently, (9.2.5) says that

$$(9.2.6) \quad N(v) \leq \|A^{-1}\|_{op} N(A(v))$$

for every $v \in \mathbf{R}^n$. Note that $\|A^{-1}\|_{op} > 0$, because $A^{-1} \neq 0$. Thus (9.2.6) is the same as saying that

$$(9.2.7) \quad \|A^{-1}\|_{op}^{-1} N(v) \leq N(A(v))$$

for every $v \in \mathbf{R}^n$.

9.2.2 Small perturbations of A

Suppose that $B \in \mathcal{L}(\mathbf{R}^n)$ satisfies

$$(9.2.8) \quad \|A - B\|_{op} < 1/\|A^{-1}\|_{op}.$$

Using (9.2.7), we get that

$$(9.2.9) \quad (\|A^{-1}\|_{op}^{-1} - \|A - B\|_{op}) N(v) \leq N(B(v))$$

for every $v \in \mathbf{R}^n$, as in Subsection 7.11.1.

In particular, this implies that $B(v) = 0$ only when $v = 0$, so that B is injective. It follows that $B(\mathbf{R}^n) = \mathbf{R}^n$, and hence

$$(9.2.10) \quad B \in GL(\mathbf{R}^n),$$

as before. This shows that

$$(9.2.11) \quad GL(\mathbf{R}^n) \text{ is an open set in } \mathcal{L}(\mathbf{R}^n),$$

with respect to the metric (9.2.2) associated to the operator norm.

One could also get (9.2.11) using determinants, as in Subsection 9.7.2.

If $A = I$, then the invertibility of B could be obtained as in Subsection 7.13.1. It is not difficult to reduce to that case, because (9.2.8) implies that

$$(9.2.12) \quad \|I - A^{-1} \circ B\|_{op} = \|A^{-1} \circ (A - B)\|_{op} \leq \|A^{-1}\|_{op} \|A - B\|_{op} < 1.$$

9.2.3 Estimating $\|B^{-1}\|_{op}$

Note that

$$(9.2.13) \quad N(B^{-1}(u)) \leq (\|A^{-1}\|_{op}^{-1} - \|A - B\|_{op})^{-1} N(u)$$

for every $u \in \mathbf{R}^n$, by (9.2.9). This means that

$$(9.2.14) \quad \begin{aligned} \|B^{-1}\|_{op} &\leq (\|A^{-1}\|_{op}^{-1} - \|A - B\|_{op})^{-1} \\ &= \frac{\|A^{-1}\|_{op}}{1 - \|A^{-1}\|_{op} \|A - B\|_{op}} \end{aligned}$$

in this situation.

Suppose that

$$(9.2.15) \quad \|A - B\|_{op} \leq 1/(2\|A^{-1}\|_{op}),$$

or equivalently

$$(9.2.16) \quad \|A^{-1}\|_{op} \|A - B\|_{op} \leq 1/2.$$

This implies that

$$(9.2.17) \quad \|B^{-1}\|_{op} \leq 2\|A^{-1}\|_{op},$$

by (9.2.14). Of course, (9.2.15) implies (9.2.8).

We would like to look at continuity properties of

$$(9.2.18) \quad A \mapsto A^{-1}$$

on $GL(\mathbf{R}^n)$. If $A, B \in GL(\mathbf{R}^n)$, then it is easy to see that

$$(9.2.19) \quad A^{-1} - B^{-1} = A^{-1} \circ (B - A) \circ B^{-1}.$$

This implies that

$$(9.2.20) \quad \|A^{-1} - B^{-1}\|_{op} \leq \|A^{-1}\|_{op} \|B^{-1}\|_{op} \|A - B\|_{op}.$$

If (9.2.15) holds, then we get that

$$(9.2.21) \quad \|A^{-1} - B^{-1}\|_{op} \leq 2 \|A^{-1}\|_{op}^2 \|A - B\|_{op},$$

by (9.2.17). It follows that $B \mapsto B^{-1}$ is continuous at A , with respect to the metric (9.2.2) associated to the operator norm, and its restriction to $GL(\mathbf{R}^n)$.

One could also get this using determinants, as in Subsection 9.7.3. This also uses the fact that $t \mapsto 1/t$ is continuous on $\mathbf{R} \setminus \{0\}$.

9.3 The inverse function theorem

Let n be a positive integer, and let N be a norm on \mathbf{R}^n , so that $d_N(v, w) = N(v - w)$ is a metric on \mathbf{R}^n . Using N , we get the corresponding operator norm $\|\cdot\|_{op}$ on the space $\mathcal{L}(\mathbf{R}^n)$ of linear mappings from \mathbf{R}^n into itself, and its associated metric, as in (9.2.2). Of course, one can simply take N to be the standard Euclidean norm on \mathbf{R}^n .

Let W be an open subset of \mathbf{R}^n , and let f be a continuously-differentiable mapping from W into \mathbf{R}^n . Also let $w \in W$ be given, and suppose that

$$(9.3.1) \quad f'(w) \text{ is invertible}$$

as a linear mapping from \mathbf{R}^n into itself. Under these conditions, the *inverse function theorem* states that there are open subsets U and V of \mathbf{R}^n with the following properties, which we describe in two parts.

9.3.1 The first part of the inverse function theorem

In the first part,

$$(9.3.2) \quad w \in U, U \subseteq W, \text{ and } f(w) \in V.$$

We are able to choose U and V so that

$$(9.3.3) \quad f \text{ is a one-to-one mapping from } U \text{ onto } V.$$

In particular, this means that $f(U) = V$ is an open set in \mathbf{R}^n .

We can also choose U so that

$$(9.3.4) \quad f'(x) \text{ is invertible as a linear mapping on } \mathbf{R}^n \text{ for each } x \in U.$$

9.3.2 The second part of the inverse function theorem

It follows from the first part that we can define a mapping g from V onto U to be the inverse of the restriction of f to U , so that

$$(9.3.5) \quad g(f(x)) = x$$

for every $x \in U$. The second part of the inverse function theorem says that

$$(9.3.6) \quad g \text{ is continuously-differentiable as a mapping from } V \text{ into } \mathbf{R}^n,$$

with

$$(9.3.7) \quad g'(y) = (f'(g(y)))^{-1}$$

for every $y \in V$. More precisely, the right side of (9.3.7) is the inverse of $f'(g(y))$, as a linear mapping on \mathbf{R}^n .

If $x \in U$ and g is differentiable at $f(x) \in V$, then the chain rule implies that

$$(9.3.8) \quad g'(f(x)) \circ f'(x) = I,$$

because of (9.3.5). This is the same as (9.3.7), with $y = f(x)$.

The proof of the inverse function theorem will be discussed in the next section. Note that the proof could be simplified when $n = 1$, using the mean value theorem and the intermediate value theorem, which is related to Exercise 2 on p114 of [210].

9.4 Proving the inverse function theorem

Let us continue with the same notation and hypotheses as in the previous section. Remember that $f'(x)$ is continuous as a function of $x \in W$ with values in $\mathcal{L}(\mathbf{R}^n)$, as in Subsection 8.10.3. We have also seen that the collection $GL(\mathbf{R}^n)$ of invertible linear mapping from \mathbf{R}^n into itself is an open set in $\mathcal{L}(\mathbf{R}^n)$, as in Subsection 9.2.2. It follows that

$$(9.4.1) \quad \{x \in W : f'(x) \in GL(\mathbf{R}^n)\}$$

is an open set in \mathbf{R}^n , because W is an open set, by hypothesis. Of course, (9.4.1) is the same as

$$(9.4.2) \quad (f')^{-1}(GL(\mathbf{R}^n)),$$

where f' is considered as a mapping from W into $\mathcal{L}(\mathbf{R}^n)$.

More precisely, the continuity of f' on W can be used to get that (9.4.1) is a relatively open set in W . This implies that (9.4.1) is an open set in \mathbf{R}^n , because W is an open set.

9.4.1 Reducing to the case where $f'(w) = I$ on \mathbf{R}^n

To prove the inverse function theorem, we can reduce to the case where

$$(9.4.3) \quad f'(w) \text{ is the identity mapping } I \text{ on } \mathbf{R}^n.$$

More precisely, this can be obtained by replacing f with

$$(9.4.4) \quad (f'(w))^{-1} \circ f,$$

as a continuously-differentiable mapping from W into \mathbf{R}^n . It is easy to see that the differential of (9.4.4) at w is equal to

$$(9.4.5) \quad (f'(w))^{-1} \circ f'(w) = I,$$

by construction.

If we can prove the inverse function theorem in this case, then the analogous conclusions for f can be obtained using

$$(9.4.6) \quad f = f'(w) \circ ((f'(w))^{-1} \circ f),$$

where $f'(w)$ is considered as a linear mapping from \mathbf{R}^n into itself.

9.4.2 The case where $f'(w) = I$

Let $B_N(w, r)$ be the open ball in \mathbf{R}^n centered at w with radius $r > 0$ with respect to $d_N(\cdot, \cdot)$, as before. Note that

$$(9.4.7) \quad B_N(w, r) \subseteq W$$

when r is sufficiently small, because W is an open set that contains w . We also have that $f'(x) - I$ is as small as we want when $x \in \mathbf{R}^n$ is close enough to w , because $f'(x)$ is continuous at w .

In particular, we can choose $r > 0$ small enough so that (9.4.7) holds and

$$(9.4.8) \quad \|f'(x) - I\|_{op} \leq 1/2$$

for every $x \in \mathbf{R}^n$ with $N(x - w) < r$. This implies that $f'(x)$ is invertible when $N(x - w) < r$, as in Subsection 9.2.2.

9.4.3 A Lipschitz condition for $f(x) - x$ on $B_N(w, r)$

Using (9.4.8), we get that

$$(9.4.9) \quad \begin{aligned} &\text{the restriction of } f(x) - x \text{ to } B_N(w, r) \\ &\text{is Lipschitz with constant } c = 1/2, \end{aligned}$$

with respect to $d_N(\cdot, \cdot)$ and its restriction to $B_N(w, r)$, as in Subsection 8.11.3. This uses the fact that $f(x) - x$ defines a continuously-differentiable mapping from W into \mathbf{R}^n , whose differential is equal to $f'(x) - I$ for each $x \in W$.

It follows that

$$(9.4.10) \quad f \text{ is bilipschitz on } B_N(w, r),$$

as in Section 7.11. In particular, f is one-to-one on $B_N(w, r)$.

We get that the restriction of f to $B_N(w, r)$ is an open mapping too, as in Subsection 7.15.2. This means that $f(B_N(w, r))$ is an open set in \mathbf{R}^n in particular.

The first part of the inverse function theorem now follows by taking

$$(9.4.11) \quad U = B_N(w, r)$$

and

$$(9.4.12) \quad V = f(B_N(w, r)).$$

9.4.4 Showing the second part

Let g be the inverse of the restriction of f to $B_N(w, r)$, as before. This is a Lipschitz mapping from V into \mathbf{R}^n , with respect to $d_N(\cdot, \cdot)$ and its restriction to V , because f is bilipschitz on $B_N(w, r)$, as in the previous subsection.

One can verify that

$$(9.4.13) \quad g \text{ is differentiable on } V,$$

with differential as in (9.3.7), using the differentiability of f . This corresponds to part of part (b) of Theorem 9.24 on p221 of [210], for instance.

The continuity of g' on V follows from (9.3.7), because f' and g are continuous, and using the continuity of (9.2.18) on $GL(\mathbf{R}^n)$.

9.5 Some remarks about \mathbf{R}^{n+m}

Let n and m be positive integers, so that \mathbf{R}^n , \mathbf{R}^m , and \mathbf{R}^{n+m} are the usual spaces of n -tuples, m -tuples, and $(n+m)$ -tuples of real numbers, respectively.

9.5.1 Identifying \mathbf{R}^{n+m} with $\mathbf{R}^n \times \mathbf{R}^m$

It will sometimes be convenient for us to identify \mathbf{R}^{n+m} with the Cartesian product $\mathbf{R}^n \times \mathbf{R}^m$ of \mathbf{R}^n and \mathbf{R}^m . Thus, if $x = (x_1, \dots, x_n) \in \mathbf{R}^n$ and $y = (y_1, \dots, y_m) \in \mathbf{R}^m$, then we may identify

$$(9.5.1) \quad (x, y) \in \mathbf{R}^n \times \mathbf{R}^m$$

with

$$(9.5.2) \quad (x_1, \dots, x_n, y_1, \dots, y_m) \in \mathbf{R}^{n+m}.$$

In particular, $(x, 0)$ and $(0, y)$ can be identified with elements of \mathbf{R}^{n+m} .

If f is a function defined on a subset of \mathbf{R}^{n+m} , then we may use

$$(9.5.3) \quad f(x, y)$$

to denote the value of f at the point in \mathbf{R}^{n+m} identified with (x, y) , when that point is in the domain of f .

9.5.2 Some related linear mappings

Let A be a linear mapping from \mathbf{R}^{n+m} into \mathbf{R}^n . If $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^m$, then put

$$(9.5.4) \quad A_1(x) = A(x, 0)$$

and

$$(9.5.5) \quad A_2(y) = A(0, y).$$

This defines A_1 and A_2 as linear mappings from \mathbf{R}^n and \mathbf{R}^m into \mathbf{R}^n , respectively. Note that

$$(9.5.6) \quad A(x, y) = A_1(x) + A_2(y)$$

for every $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^m$.

Conversely, if A_1 and A_2 are linear mappings from \mathbf{R}^n and \mathbf{R}^m into \mathbf{R}^n , respectively, then (9.5.6) defines a linear mapping from \mathbf{R}^{n+m} into \mathbf{R}^n .

9.5.3 The kernel of A

Let A be a linear mapping from \mathbf{R}^{n+m} into \mathbf{R}^n , and let A_1 and A_2 be as in (9.5.4) and (9.5.5), respectively. Also let $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^m$ be given, and observe that

$$(9.5.7) \quad A(x, y) = 0$$

if and only if

$$(9.5.8) \quad A_1(x) = -A_2(y),$$

by (9.5.6).

If A_1 is a one-to-one mapping from \mathbf{R}^n onto itself, then (9.5.8) is the same as saying that

$$(9.5.9) \quad x = -A_1^{-1}(A_2(y)).$$

In particular, for each $y \in \mathbf{R}^m$ there is a unique $x \in \mathbf{R}^n$ such that (9.5.7) holds in this situation.

9.5.4 Surjectivity of A_1 and A

Clearly

$$(9.5.10) \quad A_1(\mathbf{R}^n) \subseteq A(\mathbf{R}^{n+m}),$$

by the definition (9.5.4) of A_1 . In particular, if $A_1(\mathbf{R}^n) = \mathbf{R}^n$, then

$$(9.5.11) \quad A(\mathbf{R}^{n+m}) = \mathbf{R}^n.$$

9.5.5 An associated linear mapping \hat{A}

Consider the mapping \hat{A} from \mathbf{R}^{n+m} into itself defined by

$$(9.5.12) \quad \hat{A}(x, y) = (A(x, y), y)$$

for every $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^m$. This is a linear mapping from \mathbf{R}^{n+m} into itself, which can also be expressed as

$$(9.5.13) \quad \hat{A}(x, y) = (A_1(x) + A_2(y), y)$$

for every $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^m$, by (9.5.6).

9.5.6 Invertibility of \hat{A} and A_1

If A_1 is a one-to-one mapping from \mathbf{R}^n onto itself, then one can check that

$$(9.5.14) \quad \hat{A} \text{ is a one-to-one mapping from } \mathbf{R}^{n+m} \text{ onto itself.}$$

It is easy to see that the converse holds as well.

9.5.7 When A is surjective

Let u_1, \dots, u_{n+m} be the standard basis vectors in \mathbf{R}^{n+m} , so that for each $l = 1, \dots, m+n$, the l th coordinate of u_l is equal to 1, and the other coordinates of u_l are equal to 0. If A is a linear mapping from \mathbf{R}^{n+m} onto \mathbf{R}^n , then

$$(9.5.15) \quad \mathbf{R}^n \text{ is spanned by } A(u_1), \dots, A(u_{n+m}).$$

Under these conditions, it is well known that there is a subset K of

$$(9.5.16) \quad \{1, \dots, n+m\}$$

such that K has exactly n elements, and

$$(9.5.17) \quad A(u_k), k \in K, \text{ forms a basis for } \mathbf{R}^n.$$

If $K = \{1, \dots, n\}$, then the linear mapping A_1 on \mathbf{R}^n associated to A as in (9.5.4) is invertible. Otherwise, one could rearrange the coordinates on \mathbf{R}^{n+m} to reduce to this case.

9.6 The implicit function theorem

Let n and m be positive integers again, and let O be an open subset of \mathbf{R}^{n+m} , with respect to the standard Euclidean metric. Also let f be a continuously-differentiable mapping from O into \mathbf{R}^n . Suppose that $a \in \mathbf{R}^n$ and $b \in \mathbf{R}^m$ have the properties that $(a, b) \in O$ and

$$(9.6.1) \quad f(a, b) = 0,$$

where (a, b) is identified with an element of \mathbf{R}^{n+m} , as in the previous section.

Put
(9.6.2)
$$A = f'(a, b),$$

which is a linear mapping from \mathbf{R}^{n+m} into \mathbf{R}^n . This leads to a linear mappings A_1 and A_2 from \mathbf{R}^n and \mathbf{R}^m into \mathbf{R}^n , respectively, as in (9.5.4) and (9.5.5).

9.6.1 The statement of the implicit function theorem

Suppose that

(9.6.3) A_1 is a one-to-one mapping from \mathbf{R}^n onto itself.

Under these conditions, the *implicit function theorem* states that there are open sets $U \subseteq \mathbf{R}^{n+m}$ and $W \subseteq \mathbf{R}^m$ with the following properties, as in Theorem 9.28 on p224 of [210].

First,
(9.6.4) $(a, b) \in U, U \subseteq O, \text{ and } b \in W.$

Second, for each $y \in W$ there is a unique $x \in \mathbf{R}^n$ such that

(9.6.5) $(x, y) \in U \quad \text{and} \quad f(x, y) = 0.$

If $y \in W$, then let $g(y)$ be the element x of \mathbf{R}^n just mentioned, so that g is a mapping from W into \mathbf{R}^n , $g(b) = a$, and for every $y \in W$ we have that

(9.6.6) $(g(y), y) \in U$

and

(9.6.7) $f(g(y), y) = 0.$

We also have that g is continuously differentiable on W , with

(9.6.8) $g'(b) = -A_1^{-1} \circ A_2.$

9.6.2 Proving the implicit function theorem

To prove the implicit function theorem, we put

(9.6.9)
$$F(x, y) = (f(x, y), y)$$

for every $(x, y) \in O$, where the right side is identified with an element of \mathbf{R}^{n+m} , as before. This defines F as a continuously-differentiable mapping from O into \mathbf{R}^{n+m} , with

(9.6.10)
$$F'(a, b) = \hat{A},$$

where \hat{A} is as in (9.5.12).

In this situation, \hat{A} is a one-to-one linear mapping from \mathbf{R}^{n+m} onto itself, as in Subsection 9.5.6. Thus the inverse function theorem can be applied to F at (a, b) .

9.6.3 Using the inverse function theorem

It follows that there are open sets U, V in \mathbf{R}^{n+m} such that $(a, b) \in U, U \subseteq O$,

$$(9.6.11) \quad F(a, b) = (f(a, b), b) = (0, b) \in V,$$

and F is a one-to-one mapping from U onto V . Let G be the inverse of the restriction of F to U , so that

$$(9.6.12) \quad G(F(x, y)) = (x, y)$$

for every $(x, y) \in U$. This is a continuously-differentiable mapping from V onto U . Note that G preserves the last m coordinates of points in its domain, because of the analogous property of F .

9.6.4 The rest of the argument

It is easy to see that

$$(9.6.13) \quad W = \{y \in \mathbf{R}^m : (0, y) \in V\}$$

is an open subset of \mathbf{R}^m that contains b , because V is an open set in \mathbf{R}^{n+m} that contains $(0, b)$. If $y \in W$, then we can take $g(y) \in \mathbf{R}^n$ so that

$$(9.6.14) \quad (g(y), y) = G(0, y).$$

The continuous differentiability of g follows from the same property of G . See [210] or other texts for more details.

9.7 Some remarks about determinants

Let n be a positive integer, and let $[a_{j,l}]$ be an $n \times n$ matrix of real or complex numbers. The *determinant* $\det[a_{j,l}]$ of $[a_{j,l}]$ can be defined as a real or complex number, as appropriate, in a standard way, as mentioned in Section 7.6. More precisely, $\det[a_{j,l}]$ is a polynomial of degree n in the entries $a_{j,l}$. In particular,

$$(9.7.1) \quad \det[a_{j,l}] \text{ is continuous}$$

as a real or complex-valued function on the spaces $M_{n,n}(\mathbf{R}), M_{n,n}(\mathbf{C})$ of $n \times n$ matrices with entries in \mathbf{R}, \mathbf{C} , respectively, with respect to the metric associated to the Hilbert–Schmidt norm.

9.7.1 Determinants of linear mappings

Now let A be a linear mapping from \mathbf{R}^n or \mathbf{C}^n into itself. Remember that A corresponds to an $n \times n$ matrix of real or complex numbers, as appropriate, as in Subsection 7.2.2. The *determinant* $\det A$ of A is defined to be the determinant of the corresponding matrix, as in Section 7.6 again. This defines continuous real and complex-valued functions on the spaces $\mathcal{L}(\mathbf{R}^n), \mathcal{L}(\mathbf{C}^n)$ of linear mappings

from \mathbf{R}^n , \mathbf{C}^n into themselves, respectively, with respect to the metrics associated to the corresponding Hilbert–Schmidt norms. One could also use the metrics associated to the operator norms corresponding to the standard Euclidean norms on \mathbf{R}^n , \mathbf{C}^n , or to any other norms on \mathbf{R}^n , \mathbf{C}^n .

9.7.2 Invertibility and nonzero determinant

It is well known that a linear mapping A from \mathbf{R}^n or \mathbf{C}^n into itself is invertible if and only if

$$(9.7.2) \quad \det A \neq 0.$$

The fact that the sets of invertible elements of $\mathcal{L}(\mathbf{R}^n)$, $\mathcal{L}(\mathbf{C}^n)$ are open sets with respect to the metrics associated to the Hilbert–Schmidt or operator norms can be obtained from this, and the continuity of the determinant.

9.7.3 Continuity of $A \mapsto A^{-1}$

If A is an invertible linear mapping on \mathbf{R}^n or \mathbf{C}^n , then it is well known that the matrix corresponding to the inverse of A can be obtained from the determinant of A and the determinants of $(n-1) \times (n-1)$ submatrices of the matrix corresponding to A , as in *Cramer's rule*. This can be used to get the continuity of

$$(9.7.3) \quad A \mapsto A^{-1}$$

on the spaces of invertible elements of $\mathcal{L}(\mathbf{R}^n)$ or $\mathcal{L}(\mathbf{C}^n)$, with respect to the restrictions of the metrics associated to the Hilbert–Schmidt or operator norms to these spaces.

9.8 Local embeddings

Let n and m be positive integers, and let us continue to identify \mathbf{R}^{n+m} with $\mathbf{R}^n \times \mathbf{R}^m$, as in Subsection 9.5.1. Suppose that A is a linear mapping from \mathbf{R}^n into \mathbf{R}^{n+m} . If $v \in \mathbf{R}^n$, then $A(v) \in \mathbf{R}^{n+m}$ can be expressed as

$$(9.8.1) \quad A(v) = (A_1(v), A_2(v)),$$

where $A_1(v) \in \mathbf{R}^n$ and $A_2(v) \in \mathbf{R}^m$, using the identification just mentioned.

More precisely, this defines A_1 and A_2 as linear mappings from \mathbf{R}^n into \mathbf{R}^n and \mathbf{R}^m , respectively. Conversely if A_1 and A_2 are linear mappings from \mathbf{R}^n into \mathbf{R}^n and \mathbf{R}^m , respectively, then (9.8.1) defines a linear mapping from \mathbf{R}^n into \mathbf{R}^{n+m} .

9.8.1 Injectivity of A_1 and A

Suppose now that

$$(9.8.2) \quad A_1 \text{ is injective on } \mathbf{R}^n,$$

which implies that

$$(9.8.3) \quad A_1 \text{ is invertible on } \mathbf{R}^n,$$

as before. Of course, (9.8.2) also implies that

$$(9.8.4) \quad A \text{ is injective on } \mathbf{R}^n.$$

It can be shown that (9.8.4) implies a condition like (9.8.2), after rearranging the coordinates on \mathbf{R}^{n+m} in a suitable way. We shall discuss this further at the end of the section.

Note that

$$(9.8.5) \quad A_2 \circ A_1^{-1}$$

defines a linear mapping from \mathbf{R}^n into \mathbf{R}^m . Put

$$(9.8.6) \quad B(v) = (v, A_2(A_1^{-1}(v)))$$

for every $v \in \mathbf{R}^n$, which defines a linear mapping from \mathbf{R}^n into \mathbf{R}^{n+m} . Thus

$$(9.8.7) \quad A = B \circ A_1,$$

as a mapping from \mathbf{R}^n into \mathbf{R}^{n+m} .

9.8.2 Mappings from open sets in \mathbf{R}^n into \mathbf{R}^{n+m}

Let W be an open subset of \mathbf{R}^n , and let f be a mapping from W into \mathbf{R}^{n+m} . If $w \in W$, then $f(w) \in \mathbf{R}^{n+m}$ can be expressed as

$$(9.8.8) \quad f(w) = (f_1(w), f_2(w)),$$

where $f_1(w) \in \mathbf{R}^n$ and $f_2(w) \in \mathbf{R}^m$, as before. This defines f_1 and f_2 as mappings from W into \mathbf{R}^n and \mathbf{R}^m , respectively. Conversely, if f_1 and f_2 are mappings from W into \mathbf{R}^n and \mathbf{R}^m , respectively, then (9.8.8) defines f as a mapping from W into \mathbf{R}^{n+m} .

It is easy to see that f is differentiable at a point $w \in W$ if and only if

$$(9.8.9) \quad f_1 \text{ and } f_2 \text{ are differentiable at } w.$$

In this case,

$$(9.8.10) \quad f'(w)(v) = (f'_1(w)(v), f'_2(w)(v))$$

for every $v \in \mathbf{R}^n$. Equivalently, this means that $A = f'(w)$ corresponds to $A_1 = f'_1(w)$ and $A_2 = f'_2(w)$ as in (9.8.1).

9.8.3 Invertibility of $f'_1(w)$

Similarly, one can check that f is continuously-differentiable as a mapping from W into \mathbf{R}^{n+m} if and only if

$$(9.8.11) \quad f_1 \text{ and } f_2 \text{ are continuously-differentiable}$$

as mappings from W into \mathbf{R}^n and \mathbf{R}^m , respectively. In this situation, suppose that $w \in W$ has the property that

$$(9.8.12) \quad f'_1(w) \text{ is invertible}$$

as a linear mapping from \mathbf{R}^n into itself. Thus the inverse function theorem implies that the restriction of f_1 to a suitable neighborhood U of w is a one-to-one mapping onto a neighborhood V of $f_1(w)$, and that the corresponding inverse mapping has some additional nice properties.

This can be used to obtain more information about the behavior of f near w . Of course, this is all much simpler when f_1 is the identity mapping on \mathbf{R}^n . Otherwise, one can try to reduce to this case, at least locally near w . If f_1 has a local inverse near w , as in the preceding paragraph, then f can be expressed near w as the composition of f_1 with a simpler mapping into \mathbf{R}^{n+m} .

9.8.4 Getting an invertibility condition

Let e_1, \dots, e_n be the standard basis vectors in \mathbf{R}^n , and let u_1, \dots, u_{n+m} be the standard basis vectors for \mathbf{R}^{n+m} . If (9.8.4) holds, then

$$(9.8.13) \quad A(e_1), \dots, A(e_n) \text{ are linearly independent in } \mathbf{R}^{n+m}.$$

In this case, it is well known that there is a set $L \subseteq \{1, \dots, n+m\}$ with exactly m elements such that

$$(9.8.14) \quad A(e_1), \dots, A(e_n) \text{ together with } u_k, k \in L, \text{ is a basis for } \mathbf{R}^{n+m}.$$

If $L = \{n+1, \dots, n+m\}$, then it is easy to see that

$$(9.8.15) \quad A_1(\mathbf{R}^n) = \mathbf{R}^n,$$

so that (9.8.3) holds. Otherwise, one can get an analogous condition with the coordinates on \mathbf{R}^{n+m} rearranged.

9.9 Ranks of linear mappings

In this section, we suppose that the reader has some familiarity with linear algebra on Euclidean spaces. Let m and n be arbitrary positive integers, and let A be a linear mapping from \mathbf{R}^n into \mathbf{R}^m . It is easy to see that the image $A(\mathbf{R}^n)$ of \mathbf{R}^n under A is a linear subspace of \mathbf{R}^m . The *rank* of A is defined to be

$$(9.9.1) \quad \text{the dimension of } A(\mathbf{R}^n).$$

Thus the rank of A is a nonnegative integer less than or equal to m .

Of course, the rank of A is equal to 0 exactly when $A = 0$ on \mathbf{R}^n . The rank of A is equal to m exactly when $A(\mathbf{R}^n) = \mathbf{R}^m$.

9.9.1 Some linear subspaces of \mathbf{R}^n

Suppose that A has rank $r \geq 1$, and let w_1, \dots, w_r be a basis for $A(\mathbf{R}^n)$. If $j \in \{1, \dots, r\}$, then choose $v_j \in \mathbf{R}^n$ such that

$$(9.9.2) \quad A(v_j) = w_j.$$

Observe that

$$(9.9.3) \quad v_1, \dots, v_r \text{ are linearly independent in } \mathbf{R}^n,$$

because w_1, \dots, w_r are linearly independent in \mathbf{R}^m .

Let V be the linear span of v_1, \dots, v_r in \mathbf{R}^n , so that

$$(9.9.4) \quad V \text{ is a linear subspace of } \mathbf{R}^n \text{ of dimension } r.$$

In particular,

$$(9.9.5) \quad r \leq n$$

automatically. By construction, $A(V) = A(\mathbf{R}^n)$. One can check that

$$(9.9.6) \quad \text{the restriction of } A \text{ to } V \text{ is one-to-one}$$

in this situation.

Alternatively, let u_1, \dots, u_n be any basis for \mathbf{R}^n , such as the standard basis.

Thus

$$(9.9.7) \quad A(\mathbf{R}^n) \text{ is spanned by } A(u_1), \dots, A(u_n).$$

It is well known that a subset of the vectors $A(u_1), \dots, A(u_n)$ forms a basis for $A(\mathbf{R}^n)$. Such a subset has exactly r elements, by hypothesis. The corresponding subset of the vectors u_1, \dots, u_n can be used as a basis for a linear subspace V of \mathbf{R}^n as in the preceding paragraph.

9.9.2 Another characterization of the rank

Now let V_0 be a linear subspace of \mathbf{R}^n , and suppose that

$$(9.9.8) \quad \text{the restriction of } A \text{ to } V_0 \text{ is one-to-one.}$$

The dimension of V_0 is equal to the dimension of $A(V_0)$, which is less than or equal to the dimension of $A(\mathbf{R}^n)$. Equivalently, the dimension of V_0 is less than or equal to r .

The remarks in the previous subsection show that it is always possible to choose V_0 so that its dimension is equal to r . This implies that

$$(9.9.9) \quad \begin{array}{l} \text{the rank of } A \text{ is equal to the maximum of the dimensions} \\ \text{of the linear subspaces } V_0 \text{ of } \mathbf{R}^n \text{ on which } A \text{ is one-to-one.} \end{array}$$

If A is one-to-one on \mathbf{R}^n , then the rank of A is equal to n . Conversely, if the rank of A is equal to n , then A is one-to-one on \mathbf{R}^n .

9.9.3 Ranks of nearby linear mappings

Let V_0 be a linear subspace of \mathbf{R}^n on which A is one-to-one again, and let B be another linear mapping from \mathbf{R}^n into \mathbf{R}^m . If B is sufficiently close to A , as linear mappings from \mathbf{R}^n into \mathbf{R}^m , then

$$(9.9.10) \quad \text{the restriction of } B \text{ to } V_0 \text{ is one-to-one}$$

as well. This can be obtained from the remarks in Subsection 7.11.1.

It follows that

$$(9.9.11) \quad \text{the rank of } B \text{ is greater than or equal to the rank of } A$$

when B is sufficiently close to A . Note that the rank of B may be greater than the rank of A in this case.

9.9.4 Towards the rank theorem

Let f be a continuously-differentiable mapping from an open subset W of \mathbf{R}^n into \mathbf{R}^m , and let w be an element of W . Suppose that if $x \in W$ is sufficiently close to w , then

$$(9.9.12) \quad \text{the rank of } f'(x) \text{ is equal to the rank of } f'(w),$$

as linear mappings from \mathbf{R}^n into \mathbf{R}^m . The *rank theorem* describes the behavior of f near w in this situation, and we shall return to this in Section 9.12.

9.10 Some remarks about linear subspaces

In this section, we suppose again that the reader has some familiarity with linear algebra on Euclidean spaces, although some basic notions will also be reviewed here. In particular, if n is a positive integer, then a *linear subspace* of \mathbf{R}^n is a subset V of \mathbf{R}^n such that

$$(9.10.1) \quad u + v \in V$$

for every $u, v \in V$, and

$$(9.10.2) \quad tv \in V$$

for every $v \in V$ and $t \in \mathbf{R}$. If $V \neq \{0\}$, then it is well known that V can be spanned by finitely many of its elements. The smallest number of elements of V that span V is known as the *dimension* of V , and is denoted $\dim V$. This is a nonnegative integer less than or equal to n , which is interpreted as being equal to 0 when $V = \{0\}$.

9.10.1 Sums of linear subspaces

If V, W be linear subspaces of \mathbf{R}^n , then their *sum* may be defined as the subset of \mathbf{R}^n given by

$$(9.10.3) \quad V + W = \{v + w : v \in V, w \in W\}.$$

It is easy to see that this is a linear subspace of \mathbf{R}^n too. Note that

$$(9.10.4) \quad \dim(V + W) \leq \dim V + \dim W.$$

9.10.2 Intersections of linear subspaces

It is easy to see that the intersection of V and W is a linear subspace of \mathbf{R}^n as well. If

$$(9.10.5) \quad V \cap W = \{0\},$$

then V and W may be said to be *transverse* as linear subspaces of \mathbf{R}^n . This means that the elements of $V + W$ can be expressed in a unique way as $v + w$, with $v \in V$ and $w \in W$. In this case,

$$(9.10.6) \quad \dim(V + W) = \dim V + \dim W,$$

because one can combine bases for V and W to get a basis for $V + W$. Conversely, it is not too difficult to show that (9.10.6) implies (9.10.5).

9.10.3 Kernels of linear mappings

Let m be another positive integer, and let A be a linear mapping from \mathbf{R}^n into \mathbf{R}^m . The *kernel* or *null space* of A is defined by

$$(9.10.7) \quad \ker A = \{v \in \mathbf{R}^n : A(v) = 0\}.$$

It is easy to see that this is also a linear subspace of \mathbf{R}^n . One can check that

$$(9.10.8) \quad \ker A = \{0\}$$

if and only if A is one-to-one on \mathbf{R}^n . Of course, $A(\mathbf{R}^n)$ is a linear subspace of \mathbf{R}^m , as mentioned in Section 9.9.

9.10.4 The kernel and the rank

Suppose that $A \neq 0$, so that the dimension r of $A(\mathbf{R}^n)$ is positive. One can choose $v_1, \dots, v_r \in \mathbf{R}^n$ such that

$$(9.10.9) \quad A(v_1), \dots, A(v_r) \text{ is a basis for } A(\mathbf{R}^n),$$

as in Subsection 9.9.1. In this case, v_1, \dots, v_r are linearly independent in \mathbf{R}^n , as before. If V is the linear span of v_1, \dots, v_r in \mathbf{R}^n , then one can check that

$$(9.10.10) \quad V + \ker A = \mathbf{R}^n$$

and

$$(9.10.11) \quad V \cap (\ker A) = \{0\}.$$

This implies the well-known fact that

$$(9.10.12) \quad \dim \ker A + \dim A(\mathbf{R}^n) = n,$$

because $\dim A(\mathbf{R}^n) = r = \dim V$.

9.11 Complementary linear subspaces

Let n be a positive integer, and let V, W be linear subspaces of \mathbf{R}^n . We say that V and W are *complementary* linear subspaces of \mathbf{R}^n if

$$(9.11.1) \quad V + W = \mathbf{R}^n$$

and (9.10.5) holds. This means that every element of \mathbf{R}^n can be expressed in a unique way as a sum of elements of V and W , as before. Note that

$$(9.11.2) \quad \dim V + \dim W = n$$

in this case. If V and W are any two linear subspaces of \mathbf{R}^n that satisfy (9.11.2), then one can check that (9.10.5) is equivalent to (9.11.1).

9.11.1 Getting a complementary linear subspace

If V is any linear subspace of \mathbf{R}^n , then one can find a linear subspace W of \mathbf{R}^n such that

$$(9.11.3) \quad V \text{ and } W \text{ are complementary linear subspaces of } \mathbf{R}^n.$$

One way to do this is to start with a basis for V , and extend it to a basis for \mathbf{R}^n . In this case, one can check that the linear span of the additional basis elements is complementary to V in \mathbf{R}^n .

9.11.2 Projections

A linear mapping P from \mathbf{R}^n into itself is said to be a *projection* if

$$(9.11.4) \quad P \circ P = P$$

on \mathbf{R}^n . In this case, one can check that

$$(9.11.5) \quad \ker P \text{ and } P(\mathbf{R}^n) \text{ are complementary linear subspaces of } \mathbf{R}^n.$$

One can also verify that

$$(9.11.6) \quad I - P \text{ is a projection on } \mathbf{R}^n$$

as well. In fact,

$$(9.11.7) \quad \ker(I - P) = P(\mathbf{R}^n)$$

and

$$(9.11.8) \quad (I - P)(\mathbf{R}^n) = \ker P.$$

9.11.3 More on projections

If V and W are complementary linear subspaces of \mathbf{R}^n , then there is a unique projection P on \mathbf{R}^n such that

$$(9.11.9) \quad \ker P = V$$

and

$$(9.11.10) \quad P(\mathbf{R}^n) = W.$$

More precisely,

$$(9.11.11) \quad P(v + w) = w$$

for every $v \in V$ and $w \in W$. This can be used to define P on \mathbf{R}^n , because every element of \mathbf{R}^n can be expressed in a unique way as a sum of elements of V and W , as before.

9.12 The rank theorem

Let m , n , and r be positive integers with

$$(9.12.1) \quad r \leq \min(m, n).$$

Also let W be a nonempty open subset of \mathbf{R}^n , and let F be a continuously-differentiable mapping from W into \mathbf{R}^m . Suppose that

$$(9.12.2) \quad \text{the rank of } F'(w) \text{ is equal to } r \text{ for every } w \in W.$$

Let $a \in W$ be given, and put

$$(9.12.3) \quad A = F'(a).$$

Also put

$$(9.12.4) \quad Y_1 = A(\mathbf{R}^n),$$

which is a linear subspace of \mathbf{R}^m of dimension r . Let Y_2 be a linear subspace of \mathbf{R}^m such that

$$(9.12.5) \quad Y_1 \text{ and } Y_2 \text{ are complementary linear subspaces of } \mathbf{R}^m,$$

as in the previous section. This leads to a projection P on \mathbf{R}^m such that

$$(9.12.6) \quad P(\mathbf{R}^m) = Y_1$$

and

$$(9.12.7) \quad \ker P = Y_2,$$

as before.

Under these conditions, there are open subsets U , V of \mathbf{R}^n with the following properties. First,

$$(9.12.8) \quad a \in U \text{ and } U \subseteq W.$$

Second, there is a one-to-one mapping H from V onto U such that

$$(9.12.9) \quad H \text{ is continuously differentiable on } V,$$

$$(9.12.10) \quad H^{-1} \text{ is continuously differentiable on } U,$$

and

$$(9.12.11) \quad P(F(H(x))) = A(x)$$

for every $x \in V$. This is part of the *rank theorem*, as on p229 of [210]. More precisely, (9.12.11) corresponds to (71) on p230 of [210].

9.12.1 The case where $r = \min(m, n)$

Of course, (9.12.2) implies that

$$(9.12.12) \quad \text{the rank of } A = F'(a) \text{ is equal to } r.$$

If $w \in W$ is sufficiently close to a , then (9.12.12) implies that

$$(9.12.13) \quad \text{the rank of } F'(w) \text{ is greater than or equal to } r,$$

because $F'(w)$ is close to $F'(a)$, as in Subsection 9.9.3. Remember that the rank of any linear mapping from \mathbf{R}^n into \mathbf{R}^m is less than or equal to $\min(m, n)$, as in Section 9.9. If $r = \min(m, n)$, then (9.12.12) implies that

$$(9.12.14) \quad \begin{aligned} &\text{the rank of } F'(w) \text{ is equal to } r \\ &\text{when } w \in W \text{ is sufficiently close to } a. \end{aligned}$$

In this case, we can get (9.12.2) by replacing W with a sufficiently small open set in \mathbf{R}^n that contains a .

9.12.2 More on $r = m$ or n

If $r = m$, then (9.12.12) is the same as saying that

$$(9.12.15) \quad Y_1 = A(\mathbf{R}^n) = \mathbf{R}^m.$$

This implies that $Y_2 = \{0\}$, so that P is the identity mapping on \mathbf{R}^m . In this case, (9.12.11) means that

$$(9.12.16) \quad F(H(x)) = A(x)$$

for every $x \in V$.

If $r = m = n$, then $A = F'(a)$ is invertible on \mathbf{R}^n , and the rank theorem is basically the same as the inverse function theorem. If $r = m < n$, then the rank theorem is closely related to the implicit function theorem.

If $r = n$, then (9.12.12) is the same as saying that

$$(9.12.17) \quad A = F'(a) \text{ is injective on } \mathbf{R}^n.$$

If $r = n < m$, then the rank theorem is related to the remarks in Section 9.8.

9.13 Proving this part

Let us continue with the same notation and hypotheses as in the previous section. Note that

$$(9.13.1) \quad \dim Y_1 = r,$$

by (9.12.4) and (9.12.12). Let y_1, \dots, y_r be a basis for Y_1 .

Choose $z_j \in \mathbf{R}^n$ for each $j = 1, \dots, r$ so that

$$(9.13.2) \quad A(z_j) = y_j.$$

Let T be the unique linear mapping from Y_1 into \mathbf{R}^n such that

$$(9.13.3) \quad T(y_j) = z_j$$

for each $j = 1, \dots, r$. Thus $A(T(y_j)) = A(z_j) = y_j$ for each $j = 1, \dots, r$, so that

$$(9.13.4) \quad A(T(y)) = y$$

for every $y \in Y_1$.

If $w \in W$, then put

$$(9.13.5) \quad G(w) = w + T(P(F(w) - A(w))),$$

as in (69) on p229 of [210]. This defines a continuously-differentiable mapping from W into \mathbf{R}^n . It is easy to see that

$$(9.13.6) \quad G'(a) = I_{\mathbf{R}^n},$$

the identity mapping on \mathbf{R}^n , because $A = F'(a)$. Thus the inverse function theorem implies that there are open subsets U and V of \mathbf{R}^n such that U satisfies (9.12.8), and the restriction of G to U is a one-to-one mapping from U onto V whose inverse is continuously differentiable too. Let H be the inverse of the restriction of G to U , which is a one-to-one mapping from V onto U that satisfies (9.12.9) and (9.12.10).

Note that

$$(9.13.7) \quad H'(x) \text{ is invertible for every } x \in V,$$

as in Section 9.3. We may also suppose that

$$(9.13.8) \quad V \text{ is convex,}$$

by replacing it with a convex open subset that contains $G(a)$, if necessary, and adjusting U appropriately.

Let us check that

$$(9.13.9) \quad A \circ T \circ P \circ A = A,$$

as a linear mapping from \mathbf{R}^n into \mathbf{R}^m . More precisely,

$$(9.13.10) \quad P \circ A = A,$$

as a linear mapping from \mathbf{R}^n into $Y_1 = A(\mathbf{R}^n)$, because P is a projection of \mathbf{R}^m onto Y_1 , as in the previous section. We also have that $A \circ T$ is the identity mapping on Y_1 , as in (9.13.4). This implies (9.13.9).

If $w \in W$, then

$$(9.13.11) \quad A(G(w)) = A(w) + A(T(P(F(w) - A(w)))),$$

by (9.13.5). This implies that

$$(9.13.12) \quad A(G(w)) = A(T(P(F(w)))),$$

because of (9.13.9). It follows that

$$(9.13.13) \quad A(G(w)) = P(F(w)),$$

as in (70) on p230 of [210]. This uses (9.13.4), and the fact that P maps \mathbf{R}^m into Y_1 .

It is easy to see that (9.12.11) follows from (9.13.13), by taking $w = H(x)$, $x \in V$. Some more properties of $F \circ H$ will be discussed in the next section.

9.14 Some more properties of $F \circ H$

We continue with the same notation and hypotheses as in the previous two sections. Put

$$(9.14.1) \quad \Phi(x) = F(H(x))$$

for each $x \in V$, so that Φ is a continuously-differentiable mapping from V into \mathbf{R}^m . Of course,

$$(9.14.2) \quad \Phi'(x) = F'(H(x)) \circ H'(x)$$

for every $x \in V$, by the chain rule. This implies that

$$(9.14.3) \quad \text{the rank of } \Phi'(x) \text{ is equal to the rank of } F'(H(x))$$

for every $x \in V$, because of (9.13.7). It follows that

$$(9.14.4) \quad \text{the rank of } \Phi'(x) \text{ is equal to } r$$

for every $x \in V$, because of (9.12.2), as in (75) on p230 of [210].

Observe that

$$(9.14.5) \quad P \circ \Phi = A$$

on V , by (9.12.11). This implies that

$$(9.14.6) \quad P \circ \Phi'(x) = A$$

for every $x \in V$, by the chain rule.

Let $x \in V$ be given, and put

$$(9.14.7) \quad M = (\Phi'(x))(\mathbf{R}^n)$$

which is a linear subspace of \mathbf{R}^m . Note that

$$(9.14.8) \quad \dim M = r,$$

by (9.14.4). We also have that

$$(9.14.9) \quad P(M) = A(\mathbf{R}^n) = Y_1,$$

using (9.14.6) in the first step, and the definition (9.12.4) of Y_1 in the second step. It follows that

$$(9.14.10) \quad P \text{ is one-to-one on } M,$$

because M and Y_1 have the same dimension r .

Suppose that $h \in \mathbf{R}^n$ satisfies

$$(9.14.11) \quad A(h) = 0.$$

This means that

$$(9.14.12) \quad P((\Phi'(x))(h)) = 0,$$

because of (9.14.6). Using (9.14.10), we get that

$$(9.14.13) \quad (\Phi'(x))(h) = 0,$$

because $h \in M$.

Suppose that $y \in V$ satisfies

$$(9.14.14) \quad A(y) = A(x).$$

We would like to show that

$$(9.14.15) \quad \Phi(x) = \Phi(y),$$

which corresponds to (74) on p230 of [210]. If $t \in \mathbf{R}$ and $0 \leq t \leq 1$, then

$$(9.14.16) \quad (1-t)x + ty \in V,$$

because of (9.13.8). In this case,

$$(9.14.17) \quad \frac{d}{dt} \Phi((1-t)x + ty) = (\Phi'((1-t)x + ty))(y-x) = 0,$$

using (9.14.13) in the second step, and the fact that $A(y-x) = 0$, by (9.14.14). This implies (9.14.15).

Put

$$(9.14.18) \quad \psi(x) = \Phi(x) - A(x) = F(H(x)) - A(x)$$

for every $x \in V$, as in (72) on p230 of [210]. This defines a continuously-differentiable mapping from V into \mathbf{R}^m with

$$(9.14.19) \quad P(\psi(x)) = 0$$

for every $x \in V$, because of (9.12.11), or equivalently (9.14.5). This means that

$$(9.14.20) \quad \psi(V) \subseteq \ker P = Y_2,$$

using (9.12.7) in the second step. Note that

$$(9.14.21) \quad \psi(x) = \psi(y)$$

for all $x, y \in V$ that satisfy (9.14.14), because of (9.14.15), as in (74) on p230 of [210].

One may prefer to consider $\psi(x)$ as a function of $A(x) \in A(\mathbf{R}^n) = Y_1$, as in [210].

Chapter 10

Product spaces and related matters

10.1 Products of metric spaces

Let $(X_1, d_{X_1}), \dots, (X_n, d_{X_n})$ be finitely many metric spaces. Also let

$$(10.1.1) \quad X = X_1 \times X_2 \times \cdots \times X_n = \prod_{j=1}^n X_j$$

be the *Cartesian product* of X_1, \dots, X_n , as sets. This is the set of n -tuples $x = (x_1, \dots, x_n)$ with $x_j \in X_j$ for each $j = 1, \dots, n$.

10.1.1 Some metrics on the product

We would like to define suitable metrics on X using d_{X_1}, \dots, d_{X_n} , and indeed there are various ways to do this, as usual. If $x, y \in X$, then put

$$(10.1.2) \quad d_{X,1}(x, y) = \sum_{j=1}^n d_{X_j}(x_j, y_j).$$

It is easy to see that this defines a metric on X .

Similarly, put

$$(10.1.3) \quad d_{X,\infty}(x, y) = \max_{1 \leq j \leq n} d_{X_j}(x_j, y_j)$$

for every $x, y \in X$. One can check that this defines a metric on X as well. More precisely, the triangle inequality for $d_{X,\infty}$ follows from the triangle inequality for the norm $\|\cdot\|_\infty$ on \mathbf{R}^n , as in Subsection 1.3.2.

Put

$$(10.1.4) \quad d_{X,2}(x, y) = \left(\sum_{j=1}^n d_{X_j}(x_j, y_j)^2 \right)^{1/2}$$

for each $x, y \in X$, using of course the nonnegative square root on the right side. One can verify that this satisfies the triangle inequality, using the triangle inequality for the standard Euclidean norm on \mathbf{R}^n . It follows that $d_{X,2}$ defines another metric on X .

If $n = 1$, then $X = X_1$, and $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ are all the same as d_{X_1} .

Suppose for the moment that $X_j = \mathbf{R}$ for each $j = 1, \dots, n$, equipped with the standard Euclidean metric. In this case, X is the same as \mathbf{R}^n . The metrics $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ on X are the same as the metrics associated to the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$ on \mathbf{R}^n , respectively, as in Subsection 1.3.3.

10.1.2 Some comparisons between these metrics

Let (X_j, d_{X_j}) be any metric space for $j = 1, \dots, n$ again. Observe that

$$(10.1.5) \quad d_{X,\infty}(x, y) \leq d_{X,1}(x, y), d_{X,2}(x, y)$$

for every $x, y \in X$. We also have that

$$(10.1.6) \quad d_{X,2}(x, y) \leq d_{X,1}(x, y)$$

for every $x, y \in X$, because the standard Euclidean norm on \mathbf{R}^n is less than or equal to the norm $\|\cdot\|_1$, as in Section 1.5.

Similarly,

$$(10.1.7) \quad d_{X,2}(x, y) \leq n^{1/2} d_{X,\infty}(x, y)$$

and

$$(10.1.8) \quad d_{X,1}(x, y) \leq n d_{X,\infty}(x, y)$$

for every $x, y \in X$. As before, one can use the Cauchy–Schwarz inequality to get that

$$(10.1.9) \quad d_{X,1}(x, y) \leq n^{1/2} d_{X,2}(x, y)$$

for every $x, y \in X$.

One can use these inequalities to get that $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ have many of the same properties on X , in essentially the same way as for the analogous metrics on \mathbf{R}^n . In particular, these metrics determine the same collections of open sets, closed sets, compact sets, and so on, in X . They also determine the same limit points of subsets of X , and the same convergent sequences and Cauchy sequences. The closure of a subset of X is the same for each of these three metrics. Some of these and related properties will be discussed further in the next sections.

These metrics have many of the same properties in terms of continuity conditions for mappings between X and other metric spaces too. More precisely, the identity mapping on X is Lipschitz as a mapping from X equipped with any of these three metrics into X equipped with any other of these three metrics. This implies that the identity mapping on X is uniformly continuous and thus continuous as a mapping from X equipped with any of these three metrics into X equipped with any other of these three metrics.

10.2 Open and closed sets

Let $(X_1, d_{X_1}), \dots, (X_n, d_{X_n})$ be finitely many metric spaces again, put $X = \prod_{j=1}^n X_j$, and let $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ be the metrics defined on X as in Subsection 10.1.1. Let $x = (x_1, \dots, x_n) \in X$ and $r > 0$ be given, and for each $j = 1, \dots, n$, let

$$(10.2.1) \quad B_{X_j, d_{X_j}}(x_j, r), \quad \overline{B}_{X_j, d_{X_j}}(x_j, r)$$

be the open and closed balls in X_j centered at x_j with radius r with respect to d_{X_j} , respectively. Similarly, let

$$(10.2.2) \quad B_{X, d_{X,1}}(x, r), \quad B_{X, d_{X,2}}(x, r), \quad B_{X, d_{X,\infty}}(x, r)$$

and

$$(10.2.3) \quad \overline{B}_{X, d_{X,1}}(x, r), \quad \overline{B}_{X, d_{X,2}}(x, r), \quad \overline{B}_{X, d_{X,\infty}}(x, r)$$

be the open and closed balls in X centered at x with radius r with respect to $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$, respectively.

10.2.1 Open and closed balls in the product

It is easy to see that

$$(10.2.4) \quad B_{X, d_{X,\infty}}(x, r) = \prod_{j=1}^n B_{X_j, d_{X_j}}(x_j, r)$$

and

$$(10.2.5) \quad \overline{B}_{X, d_{X,\infty}}(x, r) = \prod_{j=1}^n \overline{B}_{X_j, d_{X_j}}(x_j, r),$$

by the definition (10.1.3) of $d_{X,\infty}$.

Using (10.1.5) and (10.1.8), we get that

$$(10.2.6) \quad B_{X, d_{X,1}}(x, r) \subseteq B_{X, d_{X,\infty}}(x, r) \subseteq B_{X, d_{X,1}}(x, nr)$$

and

$$(10.2.7) \quad \overline{B}_{X, d_{X,1}}(x, r) \subseteq \overline{B}_{X, d_{X,\infty}}(x, r) \subseteq \overline{B}_{X, d_{X,1}}(x, nr).$$

Similarly,

$$(10.2.8) \quad B_{X, d_{X,2}}(x, r) \subseteq B_{X, d_{X,\infty}}(x, r) \subseteq B_{X, d_{X,2}}(x, n^{1/2}r)$$

and

$$(10.2.9) \quad \overline{B}_{X, d_{X,2}}(x, r) \subseteq \overline{B}_{X, d_{X,\infty}}(x, r) \subseteq \overline{B}_{X, d_{X,2}}(x, n^{1/2}r),$$

by (10.1.5) and (10.1.7).

We also have that

$$(10.2.10) \quad B_{X, d_{X,1}}(x, r) \subseteq B_{X, d_{X,2}}(x, r) \subseteq B_{X, d_{X,1}}(x, n^{1/2}r)$$

and

$$(10.2.11) \quad \overline{B}_{X, d_{X,1}}(x, r) \subseteq \overline{B}_{X, d_{X,2}}(x, r) \subseteq \overline{B}_{X, d_{X,1}}(x, n^{1/2}r),$$

by (10.1.6) and (10.1.9).

10.2.2 Products of open sets

Let U_j be an open subset of X_j with respect to d_{X_j} for each $j = 1, \dots, n$. One can check that

$$(10.2.12) \quad U = \prod_{j=1}^n U_j$$

is an open set in X with respect to $d_{X,\infty}$, using (10.2.4). Indeed, if $x \in U$, then for each $j = 1, \dots, n$, $x_j \in U_j$, and there is a positive real number r_j such that

$$(10.2.13) \quad B_{X_j, d_{X_j}}(x_j, r_j) \subseteq U_j.$$

If $r = \min(r_1, \dots, r_n) > 0$, then we get that

$$(10.2.14) \quad B_{X, d_{X,\infty}}(x, r) \subseteq U.$$

It follows that U is an open set with respect to $d_{X,1}$ and $d_{X,2}$ as well, by (10.2.6) and (10.2.8).

10.2.3 Closures of products

Let $A_j \subseteq X_j$ be given for each $j = 1, \dots, n$, and put

$$(10.2.15) \quad A = \prod_{j=1}^n A_j.$$

One can verify that the closure of A in X with respect to $d_{X,\infty}$ is equal to

$$(10.2.16) \quad \prod_{j=1}^n \overline{A_j},$$

where $\overline{A_j}$ is the closure of A_j in X_j with respect to d_{X_j} for every $j = 1, \dots, n$. Basically, an element x of X is as close as we want to an element of A with respect to $d_{X,\infty}$ if and only if for each $j = 1, \dots, n$, x_j is as close as we want to an element of A_j with respect to d_{X_j} . This is the same as the closure of A in X with respect to $d_{X,1}$ and $d_{X,2}$.

10.2.4 Products of closed sets

In particular, if

$$(10.2.17) \quad A_j \text{ is a closed set in } X_j$$

with respect to d_{X_j} for each $j = 1, \dots, n$, then

$$(10.2.18) \quad A \text{ is a closed set in } X$$

with respect to $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$.

10.3 Sequences and bounded sets

Let $(X_1, d_{X_1}), \dots, (X_n, d_{X_n})$ be finitely many nonempty metric spaces, put $X = \prod_{j=1}^n X_j$, and let $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ be the metrics defined on X as in Subsection 10.1.1. Also let $\{x(l)\}_{l=1}^\infty$ be a sequence of elements of X , so that

$$(10.3.1) \quad x(l) = (x_1(l), \dots, x_n(l))$$

for each $l \geq 1$.

10.3.1 Convergent sequences in the product

One can check that

$$(10.3.2) \quad \{x(l)\}_{l=1}^\infty \text{ converges to } x = (x_1, \dots, x_n) \in X$$

with respect to $d_{X,\infty}$ if and only if

$$(10.3.3) \quad \{x_j(l)\}_{l=1}^\infty \text{ converges to } x_j \text{ in } X_j$$

with respect to d_{X_j} for each $j = 1, \dots, n$. This is equivalent to the convergence of $\{x(l)\}_{l=1}^\infty$ to x in X with respect to $d_{X,1}$, and with respect to $d_{X,2}$, as well.

10.3.2 Cauchy sequences in the product

Similarly,

$$(10.3.4) \quad \{x(l)\}_{l=1}^\infty \text{ is a Cauchy sequence in } X$$

with respect to $d_{X,\infty}$ if and only if

$$(10.3.5) \quad \{x_j(l)\}_{l=1}^\infty \text{ is a Cauchy sequence in } X_j$$

with respect to d_{X_j} for each $j = 1, \dots, n$. This is equivalent to $\{x(l)\}_{l=1}^\infty$ being a Cauchy sequence in X with respect to $d_{X,1}$, and with respect to $d_{X,2}$.

Observe that the completeness of X with respect to any of the metrics $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ are equivalent. The completeness of X with respect to these metrics is equivalent to the completeness of X_j with respect to d_{X_j} for each $j = 1, \dots, n$.

10.3.3 Bounded subsets of the product

Note that the boundedness of any subset of X with respect to any of the metrics $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$ implies the boundedness of the set with respect to the other two metrics.

Let E_j be a nonempty subset of X_j for each $j = 1, \dots, n$, and put

$$(10.3.6) \quad E = \prod_{j=1}^n E_j.$$

If

$$(10.3.7) \quad E_j \text{ is bounded in } X_j$$

with respect to d_{X_j} for each $j = 1, \dots, n$, then it is easy to see that

$$(10.3.8) \quad E \text{ is bounded in } X$$

with respect to $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$. Conversely, if E is bounded in X with respect to $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$, then E_j is bounded in X_j with respect to d_{X_j} for each $j = 1, \dots, n$.

Suppose that E_j is bounded in X_j with respect to d_{X_j} for each $j = 1, \dots, n$, and let

$$(10.3.9) \quad \text{diam}_{X_j, d_{X_j}} E_j$$

be the diameter of E_j as a subset of X_j with respect to d_{X_j} for every $j = 1, \dots, n$, as in Section 4.1. One can verify that the diameter of E as a subset of X with respect to $d_{X,\infty}$ is given by

$$(10.3.10) \quad \text{diam}_{X, d_{X,\infty}} E = \max_{1 \leq j \leq n} (\text{diam}_{X_j, d_{X_j}} E_j).$$

Similarly, the diameter of E with respect to $d_{X,1}$ is given by

$$(10.3.11) \quad \text{diam}_{X, d_{X,1}} E = \sum_{j=1}^n \text{diam}_{X_j, d_{X_j}} E_j.$$

The diameter of E with respect to $d_{X,2}$ is given by

$$(10.3.12) \quad \text{diam}_{X, d_{X,2}} E = \left(\sum_{j=1}^n (\text{diam}_{X_j, d_{X_j}} E_j)^2 \right)^{1/2}.$$

10.3.4 Total boundedness in the product

The total boundedness of any subset of X with respect to any of $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ implies the total boundedness of that set with respect to the other two metrics. If

$$(10.3.13) \quad E_j \text{ is totally bounded in } X_j$$

with respect to d_{X_j} for each $j = 1, \dots, n$, then one can check that

$$(10.3.14) \quad E \text{ is totally bounded in } X$$

with respect to $d_{X,\infty}$.

More precisely, let $r > 0$ be given, and suppose that

$$(10.3.15) \quad E_j \text{ can be covered by } L_j(r) \text{ balls of radius } r$$

in X_j for each $j = 1, \dots, n$. One can verify that E can be covered by

$$(10.3.16) \quad L(r) = \prod_{j=1}^n L_j(r)$$

balls of radius r in X with respect to $d_{X,\infty}$. This uses the products of the balls in the covering of X_j for each j .

This implies that E is totally bounded with respect to $d_{X,1}$ and $d_{X,2}$, as before. Conversely, if E is totally bounded in X with respect to $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$, then E_j is totally bounded in X_j with respect to d_{X_j} for each $j = 1, \dots, n$.

10.4 Products of compact sets

Let $(X_1, d_{X_1}), \dots, (X_n, d_{X_n})$ be finitely many metric spaces, put $X = \prod_{j=1}^n X_j$, and let $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ be the metrics defined on X as in Subsection 10.1.1. If a subset of X is compact with respect to $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$, then it is compact with respect to the other two metrics. This follows from the fact that an open subset of X with respect to $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$ is an open set with respect to the other two metrics.

Similarly, if a subset of X is sequentially compact with respect to any of the metrics $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$, then one can check directly that it is sequentially compact with respect to the other two metrics. This uses the fact that convergence of sequences of elements of X with respect to $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ are the same, as in Subsections 10.1.2 and 10.3.1.

10.4.1 Compactness of products

Let K_j be a compact subset of X_j with respect to d_{X_j} for each $j = 1, \dots, n$, and put

$$(10.4.1) \quad K = \prod_{j=1}^n K_j.$$

It is well known that

$$(10.4.2) \quad K \text{ is compact in } X,$$

with respect to any of the metrics $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$. This may be considered as a particular case of a famous theorem of Tychonoff for arbitrary topological spaces, instead of metric spaces. Let us mention a couple of other ways to look at this.

10.4.2 Sequential compactness of products

If K_j is sequentially compact in X_j with respect to d_{X_j} for each $j = 1, \dots, n$, then

$$(10.4.3) \quad K \text{ is sequentially compact in } X$$

with respect to any of the metrics $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$. This can be seen using an argument like one in Subsection 5.1.2, when the set E considered there is $\{1, \dots, n\}$. If $\{x(l)\}_{l=1}^\infty$ is any sequence of elements of K , then one can pass to a subsequence such that the corresponding subsequence of $\{x_1(l)\}_{l=1}^\infty$ converges to an element of K_1 , because K_1 is sequentially compact. One can repeat the

process, using subsequences of the previous subsequences, to get a subsequence of $\{x(l)\}_{l=1}^{\infty}$ that converges to an element of K , as before. Remember that compactness and sequential compactness are equivalent in metric spaces, as in Sections 4.5 and Subsection 4.7.1.

10.4.3 Using completeness and total boundedness

Alternatively, if

$$(10.4.4) \quad K_j \text{ is compact in } X_j$$

with respect to d_{X_j} for each $j = 1, \dots, n$, then

$$(10.4.5) \quad K_j \text{ is closed and totally bounded in } X_j$$

with respect to d_{X_j} for every $j = 1, \dots, n$. This implies that

$$(10.4.6) \quad K \text{ is closed and totally bounded in } X$$

with respect to each of $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$, as in Subsections 10.2.4 and 10.3.4.

If X is complete with respect to $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$, then it follows that K is compact, as in Section 4.9. If

$$(10.4.7) \quad X_j \text{ is complete with respect to } d_{X_j}$$

for each $j = 1, \dots, n$, then X is complete with respect to each of $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$, as in Subsection 10.3.2.

Note that one can reduce to the case where $K_j = X_j$ for each $j = 1, \dots, n$, by standard results.

It is well known that compact metric spaces are complete, as in Subsection 4.10.1. Thus one can use the previous argument when $K_j = X_j$ for each $j = 1, \dots, n$.

10.4.4 Necessity of the compactness of the K_j 's

If $K_j \neq \emptyset$ for each $j = 1, \dots, n$, and K is compact in X with respect to $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$, then it is not too difficult to show that K_j is compact in X_j with respect to d_{X_j} for each $j = 1, \dots, n$. Indeed, if $x_l \in K_l$ for $l \neq j$, then it suffices to show that

$$(10.4.8) \quad \{(x_1, \dots, x_n) : x_j \in K_j\}$$

is compact in X with respect to $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$. This is related to a remark in Subsection 10.6.4, when $n = 2$.

This set (10.4.8) is the same as the Cartesian product of K_j with the one-element sets $\{x_l\}$ for $l \neq j$, and it is also the same as the intersection of K with

$$(10.4.9) \quad \{(x_1, \dots, x_n) : x_j \in X_j\}.$$

It is well known that the intersection of a compact set and a closed set is compact in any metric space.

10.5 Some mappings on products

Let $(X_1, d_{X_1}), \dots, (X_n, d_{X_n})$ be finitely many nonempty metric spaces, put $X = \prod_{j=1}^n X_j$, and let $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ be the metrics defined on X as in Subsection 10.1.1. If $x = (x_1, \dots, x_n) \in X$ and $1 \leq l \leq n$, then put

$$(10.5.1) \quad p_l(x) = x_l.$$

This defines a mapping p_l from X onto X_l .

Clearly

$$(10.5.2) \quad \begin{aligned} d_{X_l}(p_l(x), p_l(w)) = d_{X_l}(x_l, w_l) &\leq d_{X,\infty}(x, w) \\ &\leq d_{X,1}(x, w), d_{X,2}(x, w) \end{aligned}$$

for every $x, w \in X$. This means that p_l is Lipschitz with constant 1 with respect to $d_{X,\infty}$ on X , and thus with respect to $d_{X,1}$ and $d_{X,2}$ on X . In particular, p_l is uniformly continuous on X with respect to $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$.

10.5.1 Coordinate projections are open mappings

If $x \in X$ and $r > 0$, then

$$(10.5.3) \quad p_l(B_{X,d_{X,\infty}}(x, r)) = p_l\left(\prod_{j=1}^n B_{X_j,d_{X_j}}(x_j, r)\right) = B_{X_l,d_{X_l}}(x_l, r),$$

using (10.2.4) in the first step. It follows that

$$(10.5.4) \quad p_l(B_{X,d_{X,1}}(x, r)) \supseteq p_l(B_{X,d_{X,\infty}}(x, r/n)) = B_{X_l,d_{X_l}}(x_l, r/n),$$

using (10.2.6) in the first step. Similarly,

$$(10.5.5) \quad p_l(B_{X,d_{X,2}}(x, r)) \supseteq p_l(B_{X,d_{X,\infty}}(x, r/n^{1/2})) = B_{X_l,d_{X_l}}(x_l, r/n^{1/2}),$$

using (10.2.8) in the first step. This shows that p_l is an open mapping from X onto X_l , with respect to $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$ on X , as in Section 7.15.

10.5.2 Some more partial Lipschitz conditions

Let (Y, d_Y) be another metric space, let E be a subset of X , and let f be a mapping from E into Y . Let us say that f is *partially Lipschitz in the l th variable with constant $C_l \geq 0$* on E if

$$(10.5.6) \quad d_Y(f(x), f(x')) \leq C_l d_{X_l}(x_l, x'_l)$$

for every $x, x' \in E$ such that $x_j = x'_j$ when $j \neq l$. This is analogous to the condition discussed in Section 8.3 for functions defined on subsets of \mathbf{R}^n .

10.5.3 Functions on product sets

Suppose that $E = \prod_{j=1}^n E_j$ for some $E_j \subseteq X_j$, $1 \leq j \leq n$, and that f is partially Lipschitz in the l th variable with constant $C_l \geq 0$ on E for each $l = 1, \dots, n$. If $x, w \in E$, then

$$(10.5.7) \quad d_Y(f(x), f(w)) \leq \sum_{l=1}^n C_l d_{X_l}(x_l, w_l),$$

as in Section 8.3. This implies that

$$(10.5.8) \quad d_Y(f(x), f(w)) \leq \left(\max_{1 \leq l \leq n} C_l \right) d_{X,1}(x, w),$$

as before. Similarly,

$$(10.5.9) \quad d_Y(f(x), f(w)) \leq \left(\sum_{l=1}^n C_l^2 \right)^{1/2} d_{X,2}(x, w)$$

and

$$(10.5.10) \quad d_Y(f(x), f(w)) \leq \left(\sum_{l=1}^n C_l \right) d_{X,\infty}(x, w).$$

This shows that f is Lipschitz on E with respect to the restrictions of $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ to E under these conditions.

10.5.4 Joint and separate continuity

Now let f be any mapping from X into Y . If f is continuous at a point $x \in X$ with respect to any of the metrics $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$ on X , then it is easy to see that f is continuous at x with respect to the other two metrics. This is often called *joint continuity* of f at x .

It is often convenient to consider f as a function of n variables in X_1, \dots, X_n . If $1 \leq l \leq n$, then we may consider f as a function of the l th variable on X_l with values in Y , with the j th variable equal to x_j when $j \neq l$. If f is continuous as a function of the l th variable on X_l at x_l in this way for each $l = 1, \dots, n$, then f is said to be *separately continuous* at x . Note that joint continuity at x implies separate continuity.

10.6 Uniform continuity

Let (X, d_X) and (Y, d_Y) be metric spaces, let f be a mapping from X into Y , and let A be a subset of X . Let us say that f is *uniformly continuous along* A if for every $\epsilon > 0$ there is a $\delta > 0$ such that for every $x \in A$ and $w \in X$ with $d_X(x, w) < \delta$, we have that

$$(10.6.1) \quad d_Y(f(x), f(w)) < \epsilon.$$

This implies that the restriction of f to A is uniformly continuous, with respect to the restriction of d_X to A . This also implies that f is continuous at every element of A , as a mapping from X into Y . If $A = X$, then uniform continuity along A is the same as uniform continuity on X .

10.6.1 Uniform continuity along compact sets

Suppose for the moment that A is a compact subset of X . If f is continuous at every element of A , as a function on X , then one can show that f is uniformly continuous along A . This uses the same type of arguments as used to show that continuous mappings on compact metric spaces are uniformly continuous.

10.6.2 Uniform continuity and product spaces

Let (X_1, d_{X_1}) , (X_2, d_{X_2}) be metric spaces, and let us now take

$$(10.6.2) \quad X = X_1 \times X_2.$$

We can define the metrics $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ on X using d_{X_1} and d_{X_2} as in Subsection 10.1.1, with $n = 2$. If a mapping f from X into Y is uniformly continuous along a set $A \subseteq X$ with respect to any of these three metrics on X , then it is easy to see that f is uniformly continuous along A with respect to the other two metrics.

10.6.3 Uniform continuity and equicontinuity

Let $x_1 \in X_1$ be given, and consider

$$(10.6.3) \quad A = \{x_1\} \times X_2.$$

Let f be a mapping from X into Y again, and for each $x_2 \in X_2$, let us consider $f(\cdot, x_2)$ as a mapping from X_1 into Y . Let

$$(10.6.4) \quad \mathcal{E} = \{f(\cdot, x_2) : x_2 \in X_2\}$$

be the collection of these mappings from X_1 into Y .

Observe that \mathcal{E} is equicontinuous at x_1 , as a collection of mappings from X_1 into Y , if and only if for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$(10.6.5) \quad d_Y(f(x_1, x_2), f(w_1, x_2)) < \epsilon$$

for every $w_1 \in X_1$ with

$$(10.6.6) \quad d_{X_1}(x_1, w_1) < \delta,$$

and every $x_2 \in X_2$, as in Section 5.2.

If f is uniformly continuous along A with respect to any of the metrics $d_{X,1}$, $d_{X,2}$, or $d_{X,3}$ on X , then one can check that \mathcal{E} is equicontinuous at x_1 , as a collection of mappings from X_1 into Y . As before, uniform continuity along A

implies that the restriction of f to A is uniformly continuous, which in this case means that $f(x_1, \cdot)$ is uniformly continuous as a mapping from X_2 into Y .

Conversely, suppose that $f(x_1, \cdot)$ is uniformly continuous as a mapping from X_2 into Y , and that \mathcal{E} is equicontinuous at x_1 as a collection of mappings from X_1 into Y . Under these conditions, one can verify that f is uniformly continuous along A with respect to $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ on X .

10.6.4 Some compact subsets of a product space

Observe that

$$(10.6.7) \quad x_2 \mapsto (x_1, x_2)$$

defines an isometry from X_2 into X , with respect to $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ on X , as in Subsection 7.8.1. In particular, if X_2 is compact, then it follows that A is compact in X , with respect to $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$. In this case, if f is continuous at every point in A as a mapping from X into Y , then f is uniformly continuous along A , as before.

10.7 Continuity and integration

Let (X_1, d_{X_1}) be a metric space, and let a, b be real numbers, with $a < b$. Let us take

$$(10.7.1) \quad X_2 = [a, b],$$

and d_{X_2} to be the restriction of the standard Euclidean metric on \mathbf{R} to X_2 . As in the previous section, we take

$$(10.7.2) \quad X = X_1 \times X_2 = X_1 \times [a, b],$$

and $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ to be the metrics defined on X as in Subsection 10.1.1, with $n = 2$.

10.7.1 Integration in x_2

Let f be a continuous real-valued function on X , with respect to any of $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$, and thus with respect to each of these metrics. If $x_1 \in X_1$, then $f(x_1, \cdot)$ is a continuous real-valued function on X_2 , and we put

$$(10.7.3) \quad F(x_1) = \int_a^b f(x_1, x_2) dx_2.$$

This defines a real-valued function on X_1 . Note that

$$(10.7.4) \quad |F(x_1)| \leq \int_a^b |f(x_1, x_2)| dx_2 \leq (b-a) \sup_{a \leq x_2 \leq b} |f(x_1, x_2)|.$$

More precisely, the supremum on the right is finite, because $[a, b]$ is compact as a subset of the real line, and thus as a subset of itself, with respect to the standard Euclidean metric on \mathbf{R} and its restriction to $[a, b]$.

10.7.2 Continuity of F

If $x_1 \in X_1$, then

$$(10.7.5) \quad \{x_1\} \times [a, b] = \{x_1\} \times X_2$$

is compact in X , with respect to $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$, as in Subsection 10.6.4. This implies that f is uniformly continuous along (10.7.5), as in Subsection 10.6.1. This means that

$$(10.7.6) \quad \mathcal{E} = \{f(\cdot, x_2) : a \leq x_2 \leq b\}$$

is equicontinuous at x_1 as a collection of real-valued functions on X_1 , as in Subsection 10.6.3.

One can use this to check that F is continuous at x_1 , as a real-valued function on X_1 . Indeed, if $w_1 \in X_1$, then

$$(10.7.7) \quad |F(x_1) - F(w_1)| = \left| \int_a^b (f(x_1, x_2) - f(w_1, x_2)) dx_2 \right| \\ \leq \int_a^b |f(x_1, x_2) - f(w_1, x_2)| dx_2 \\ \leq (b-a) \sup_{a \leq x_2 \leq b} |f(x_1, x_2) - f(w_1, x_2)|,$$

as before. The right side can be made arbitrarily small by taking w_1 sufficiently close to x_1 in X_1 , because of the equicontinuity of \mathcal{E} at x_1 .

10.7.3 Another argument using sequences

Alternatively, let $\{x_{1,j}\}_{j=1}^\infty$ be a sequence of elements of X_1 that converges to x_1 . It is easy to see that $f(x_{1,j}, \cdot)$ converges to $f(x_1, \cdot)$ uniformly as a sequence of real-valued functions on $[a, b]$, because of the equicontinuity of \mathcal{E} at x_1 . This implies that

$$(10.7.8) \quad \lim_{j \rightarrow \infty} F(x_{1,j}) = F(x_1),$$

as in Subsection 3.4.3. It follows that F is continuous at x_1 .

10.7.4 Uniform continuity of F

If \mathcal{E} is uniformly equicontinuous on X_1 , then it is easy to see that F is uniformly continuous on X_1 , using (10.7.7). In particular, this holds when X_1 is compact, because \mathcal{E} is equicontinuous at every point in X_1 , as in Subsection 5.2.1.

If X_1 is compact, then X is compact with respect to any of the metrics $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$, as in Subsection 10.4.1, because X_2 is compact. In this case, f is uniformly continuous with respect to any of these three metrics on X . Of course, uniform continuity of f on X with respect to any of these three metrics implies that \mathcal{E} is uniformly equicontinuous on X_1 .

10.8 Iterated integrals

Let n be a positive integer, and let a_1, \dots, a_n and b_1, \dots, b_n be real numbers, with

$$(10.8.1) \quad a_j < b_j$$

for each $j = 1, \dots, n$. Let us consider

$$(10.8.2) \quad X_j = [a_j, b_j]$$

as a metric space for each $j = 1, \dots, n$, with d_{X_j} equal to the restriction of the standard Euclidean metric on \mathbf{R} to X_j .

10.8.1 The product of the first l intervals

Put

$$(10.8.3) \quad X^l = \prod_{j=1}^l X_j$$

for each $l = 1, \dots, n$. We can define metrics $d_{X^l,1}$, $d_{X^l,2}$, and $d_{X^l,\infty}$ on X^l , as in Subsection 10.1.1. We shall refer to functions on X^l as being continuous if they are continuous with respect to any of these three metrics, and thus with respect to the other two metrics.

Of course, X^l is a subset of \mathbf{R}^l for each $l = 1, \dots, n$. Note that $d_{X^l,1}$, $d_{X^l,2}$, and $d_{X^l,\infty}$ are the same as the restrictions to X^l of the metrics on \mathbf{R}^l associated to the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$, respectively, as in Subsection 1.3.3.

10.8.2 Integrating continuous functions one variable at a time

Let f be a continuous real-valued function on X^n , and put $f_n = f$. Suppose that f_l has been defined as a continuous real-valued function on X^l for some $l = 1, \dots, n$. If $l \geq 2$ and $x_j \in X_j$ for $j = 1, \dots, l-1$, then put

$$(10.8.4) \quad f_{l-1}(x_1, \dots, x_{l-1}) = \int_{a_l}^{b_l} f_l(x_1, \dots, x_{l-1}, x_l) dx_l.$$

This defines f_{l-1} as a continuous real-valued function on X^{l-1} , as in the previous section. More precisely, this uses the obvious identification of X^l with $X^{l-1} \times X_l$.

Continuing in this way, we define f_l on X^l for each $l = 1, \dots, n$. Similarly, put

$$(10.8.5) \quad f_0 = \int_{a_1}^{b_1} f_1(x_1) dx_1,$$

which is the analogue of (10.8.4) with $l = 1$, and which is simply a real number. This can be used to define the n -fold *iterated integral* of f over X^n , as on p246 of [210].

10.8.3 More on iterated integrals

Alternatively, one can define the integral of f over X^n as an n -dimensional Riemann integral. Any reasonable approach to this will give the same answer, because f is uniformly continuous on X^n , since X^n is compact.

Of course, one could also define n -fold iterated integrals of f over X^n by integrating the variables in a different order, and one would like to verify that this leads to the same result. One way to do this is to show that the iterated integrals are all the same as the corresponding n -dimensional Riemann integral. Another approach is given in Theorem 10.2 on p246 of [210].

Some arguments for interchanging the order of integration in two variables will be discussed in the next sections. This can be used repeatedly to permute the order of integration in any number of variables.

10.9 Partitions of intervals

Let a, b be real numbers, with $a < b$. Suppose that $\mathcal{P} = \{t_j\}_{j=0}^k$ is a *partition* of $[a, b]$, which is to say a finite sequence of real numbers such that

$$(10.9.1) \quad a = t_0 < t_1 < \cdots < t_{k-1} < t_k = b.$$

10.9.1 Piecewise-linear functions on $[a, b]$

A real-valued function on $[a, b]$ is said to be *piecewise linear with breakpoints in \mathcal{P}* if the function is linear on $[t_{j-1}, t_j]$ for each $j = 1, \dots, k$. It is easy to see that such a function is continuous, with respect to the standard Euclidean metric on \mathbf{R} , and its restriction to $[a, b]$.

Note that such a function is uniquely determined by its values at the points in \mathcal{P} . The values of such a function at the points in \mathcal{P} can be arbitrary real numbers. This is because linear functions on \mathbf{R} are uniquely determined by their values at any two distinct points, and those values can be arbitrary real numbers.

10.9.2 Some piecewise-linear approximations

If f is any real-valued function on $[a, b]$, then we let $A_{\mathcal{P}}(f)$ be the unique piecewise-linear function on $[a, b]$ with breakpoints in \mathcal{P} that is equal to f at the points in \mathcal{P} . If $x \in [a, b]$ and $t_{j-1} \leq x \leq t_j$ for some $j = 1, \dots, k$, then

$$(10.9.2) \quad (A_{\mathcal{P}}(f))(x) = f(t_{j-1}) (t_j - t_{j-1})^{-1} (t_j - x) + f(t_j) (t_j - t_{j-1})^{-1} (x - t_{j-1}).$$

In this case,

$$(10.9.3) \quad f(x) - (A_{\mathcal{P}}(f))(x) = (f(x) - f(t_{j-1})) (t_j - t_{j-1})^{-1} (t_j - x) + (f(x) - f(t_j)) (t_j - t_{j-1})^{-1} (x - t_{j-1}).$$

This implies that

$$(10.9.4) \quad \begin{aligned} |f(x) - (A_{\mathcal{P}}(f))(x)| &\leq |f(x) - f(t_{j-1})| (t_j - t_{j-1})^{-1} (t_j - x) \\ &\quad + |f(x) - f(t_j)| (t_j - t_{j-1})^{-1} (x - t_{j-1}). \end{aligned}$$

It follows that

$$(10.9.5) \quad |f(x) - (A_{\mathcal{P}}(f))(x)| \leq \max(|f(x) - f(t_{j-1})|, |f(x) - f(t_j)|),$$

because

$$(10.9.6) \quad \begin{aligned} (t_j - t_{j-1})^{-1} (t_j - x) + (t_j - t_{j-1})^{-1} (x - t_{j-1}) \\ = (t_j - t_{j-1})^{-1} (t_j - t_{j-1}) = 1, \end{aligned}$$

where both terms on the first line are greater than or equal to 0.

10.9.3 Approximating the integral of f

Observe that

$$(10.9.7) \quad \begin{aligned} \int_a^b (A_{\mathcal{P}}(f))(x) dx &= \sum_{j=1}^k \int_{t_{j-1}}^{t_j} (A_{\mathcal{P}}(f))(x) dx \\ &= \sum_{j=1}^k (1/2) (f(t_{j-1}) + f(t_j)) (t_j - t_{j-1}). \end{aligned}$$

Suppose that f is continuous on $[a, b]$, with respect to the standard Euclidean metric on \mathbf{R} and its restriction to $[a, b]$. Under these conditions, we have that

$$(10.9.8) \quad \begin{aligned} \left| \int_a^b f(x) dx - \int_a^b (A_{\mathcal{P}}(f))(x) dx \right| &= \left| \int_a^b (f(x) - (A_{\mathcal{P}}(f))(x)) dx \right| \\ &\leq \int_a^b |f(x) - (A_{\mathcal{P}}(f))(x)| dx \\ &\leq (b - a) \sup_{a \leq x \leq b} |f(x) - (A_{\mathcal{P}}(f))(x)|. \end{aligned}$$

10.9.4 Using uniform continuity of f

Note that f is uniformly continuous on $[a, b]$, because $[a, b]$ is compact. Using this and (10.9.5), we get that f is uniformly approximated by $A_{\mathcal{P}}(f)$ on $[a, b]$ when the partition \mathcal{P} of $[a, b]$ is sufficiently fine. More precisely, this means that

$$(10.9.9) \quad \sup_{a \leq x \leq b} |f(x) - (A_{\mathcal{P}}(f))(x)|$$

is as small as we want when

$$(10.9.10) \quad \max_{1 \leq j \leq k} (t_j - t_{j-1})$$

is sufficiently small. It follows that the left side of (10.9.8) is as small as we want when (10.9.10) is sufficiently small.

Similarly, one can check that the left side of (10.9.8) is arbitrarily small for suitable partitions \mathcal{P} of $[a, b]$ when f is Riemann integrable on $[a, b]$.

10.10 Partitions and product spaces

Let (X_1, d_{X_1}) be a metric space, and let a_2, b_2 be real numbers, with $a_2 < b_2$. Let us take

$$(10.10.1) \quad X_2 = [a_2, b_2],$$

and d_{X_2} to be the restriction of the standard Euclidean metric on \mathbf{R} to X_2 . We also take

$$(10.10.2) \quad X = X_1 \times X_2 = X_1 \times [a_2, b_2],$$

and $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ to be the metrics defined on X as in Subsection 10.1.1, with $n = 2$. We shall refer to functions on X as being continuous if they are continuous with respect to any of these three metrics, and thus with respect to the other two metrics, as before.

10.10.1 A partition of $[a_2, b_2]$

Let $\mathcal{P}_2 = \{t_{2,j}\}_{j=0}^k$ be a partition of $[a_2, b_2]$, so that

$$(10.10.3) \quad a_2 = t_{2,0} < \cdots < t_{2,k} = b_2.$$

Let f be a continuous real-valued function on X , and let us define $A_{2,\mathcal{P}_2}(f)$ as a real-valued function on X in essentially the same way as in Subsection 10.9.2, as a function of the second variable. Thus, if $x_1 \in X_1$, $x_2 \in [a_2, b_2]$, and $t_{2,j-1} \leq x_2 \leq t_{2,j}$ for some $j = 1, \dots, k$, then we put

$$(10.10.4) \quad \begin{aligned} (A_{2,\mathcal{P}_2}(f))(x_1, x_2) &= f(x_1, t_{2,j-1}) (t_{2,j} - t_{2,j-1})^{-1} (t_{2,j} - x_2) \\ &\quad + f(x_1, t_{2,j}) (t_{2,j} - t_{2,j-1})^{-1} (x_2 - t_{2,j-1}). \end{aligned}$$

Under these conditions, we get that

$$(10.10.5) \quad \begin{aligned} &|f(x_1, x_2) - (A_{2,\mathcal{P}_2}(f))(x_1, x_2)| \\ &\leq \max(|f(x_1, x_2) - f(x_1, t_{2,j-1})|, |f(x_1, x_2) - f(x_1, t_{2,j})|), \end{aligned}$$

as in (10.9.5).

10.10.2 Integrals in x_2

As in Subsection 10.7.2,

$$(10.10.6) \quad \int_{a_2}^{b_2} f(x_1, x_2) dx_2$$

defines a continuous real-valued function of $x_1 \in X_1$. The analogous statement for $A_{2,\mathcal{P}_2}(f)$ can be seen more directly. Namely,

$$(10.10.7) \quad \int_{a_2}^{b_2} (A_{2,\mathcal{P}_2}(f))(x_1, x_2) dx_2 \\ = \sum_{j=1}^k (1/2) (f(x_1, t_{2,j-1}) + f(x_1, t_{2,j})) (t_{2,j} - t_{2,j-1})$$

for every $x_1 \in X_1$, as in (10.9.7).

We also have that

$$(10.10.8) \quad \left| \int_{a_2}^{b_2} f(x_1, x_2) dx_2 - \int_{a_2}^{b_2} (A_{2,\mathcal{P}_2}(f))(x_1, x_2) dx_2 \right| \\ \leq \int_{a_2}^{b_2} |f(x_1, x_2) - (A_{2,\mathcal{P}_2}(f))(x_1, x_2)| dx_2 \\ \leq (b_2 - a_2) \sup_{a_2 \leq x_2 \leq b_2} |f(x_1, x_2) - (A_{2,\mathcal{P}_2}(f))(x_1, x_2)|$$

for every $x_1 \in X_1$, as in (10.9.8).

10.10.3 A uniform equicontinuity condition

Suppose that

$$(10.10.9) \quad \mathcal{E}_2 = \{f(x_1, \cdot) : x_1 \in X_1\}$$

is uniformly equicontinuous as a collection of real-valued functions on $[a_2, b_2]$. This implies that

$$(10.10.10) \quad \sup_{a_2 \leq x_2 \leq b_2} |f(x_1, x_2) - (A_{2,\mathcal{P}_2}(f))(x_1, x_2)|$$

is as small as we want, uniformly over $x_1 \in X_1$, when

$$(10.10.11) \quad \max_{1 \leq j \leq k} (t_{2,j} - t_{2,j-1})$$

is sufficiently small, because of (10.10.5). It follows that (10.10.6) can be uniformly approximated by (10.10.7), as real-valued functions of $x_1 \in X_1$, when (10.10.11) is sufficiently small, by (10.10.8).

If X_1 is compact, then \mathcal{E}_2 is equicontinuous at every point in $[a_2, b_2]$, as in Section 10.6. This implies that \mathcal{E}_2 is uniformly equicontinuous on $[a_2, b_2]$, as in Subsection 5.2.1, because $[a_2, b_2]$ is compact.

Alternatively, if X_1 is compact, then X is compact with respect to any of the metrics $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$, as in Subsection 10.4.1, because $[a_2, b_2]$ is compact. This means that f is uniformly continuous on X , with respect to any of these three metrics. It is easy to see that this implies that \mathcal{E}_2 is uniformly equicontinuous on $[a_2, b_2]$.

10.11 Partitions and integrals

Let a_1, b_1 be real numbers, with $a_1 < b_1$. We would like to continue with the same notation and hypotheses as in the previous section, with

$$(10.11.1) \quad X_1 = [a_1, b_1],$$

and with d_{X_1} equal to the restriction of the standard Euclidean metric on \mathbf{R} to X_1 . Thus

$$(10.11.2) \quad X = X_1 \times X_2 = [a_1, b_1] \times [a_2, b_2],$$

and $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ are the same as the restrictions to X of the metrics on \mathbf{R}^2 associated to the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$, respectively, as in Subsection 1.3.3. Note that f is uniformly continuous on X , because X is compact.

10.11.1 Integrals in x_1

We can integrate $f(x_1, x_2)$ in x_1 to get a continuous real-valued function

$$(10.11.3) \quad \int_{a_1}^{b_1} f(x_1, x_2) dx_1$$

of x_2 on $[a_2, b_2]$, as in Subsection 10.7.2. The analogous statement for $A_{2,\mathcal{P}_2}(f)$ can be seen more directly, as before. Indeed, if $x_2 \in [a_2, b_2]$ and

$$(10.11.4) \quad t_{2,j-1} \leq x_2 \leq t_{2,j}$$

for some $j = 1, \dots, k$, then

$$(10.11.5) \quad \begin{aligned} & \int_{a_1}^{b_1} (A_{2,\mathcal{P}_2}(f))(x_1, x_2) dx_1 \\ &= (t_{2,j} - t_{2,j-1})^{-1} (t_{2,j} - x_2) \int_{a_1}^{b_1} f(x_1, t_{2,j-1}) dx_1 \\ & \quad + (t_{2,j} - t_{2,j-1})^{-1} (x_2 - t_{2,j-1}) \int_{a_1}^{b_1} f(x_1, t_{2,j}) dx_1, \end{aligned}$$

by (10.10.4). Of course, this is equal to (10.11.3) when $x_2 = t_{2,j-1}$ or $t_{2,j}$.

10.11.2 Approximating the integral of f in x_1

Observe that

$$(10.11.6) \quad \begin{aligned} & \left| \int_{a_1}^{b_1} f(x_1, x_2) dx_1 - \int_{a_1}^{b_1} (A_{2,\mathcal{P}_2}(f))(x_1, x_2) dx_1 \right| \\ &= \left| \int_{a_1}^{b_1} (f(x_1, x_2) - (A_{2,\mathcal{P}_2}(f))(x_1, x_2)) dx_1 \right| \\ &\leq \int_{a_1}^{b_1} |f(x_1, x_2) - (A_{2,\mathcal{P}_2}(f))(x_1, x_2)| dx_1 \\ &\leq (b_1 - a_1) \sup_{a_1 \leq x_1 \leq b_1} |f(x_1, x_2) - (A_{2,\mathcal{P}_2}(f))(x_1, x_2)| \end{aligned}$$

for every $x_2 \in [a_2, b_2]$. We also have that

$$(10.11.7) \quad \sup_{a_1 \leq x_1 \leq b_1} |f(x_1, x_2) - (A_{2, \mathcal{P}_2}(f))(x_1, x_2)|$$

is as small as we want, uniformly over $x_2 \in [a_2, b_2]$, when (10.10.11) is sufficiently small. Equivalently, this means that

$$(10.11.8) \quad \sup_{a_1 \leq x_1 \leq b_1} \sup_{a_2 \leq x_2 \leq b_2} |f(x_1, x_2) - (A_{2, \mathcal{P}_2}(f))(x_1, x_2)|$$

is as small as we want when (10.10.11) is sufficiently small. This is the same as the analogous statement for (10.10.10), which follows from the uniform continuity of f on X , as before. Combining this with (10.11.6), we get that (10.11.3) can be uniformly approximated by (10.11.5), as real-valued functions of $x_2 \in [a_2, b_2]$, when (10.10.11) is sufficiently small.

10.11.3 Some iterated integrals

One can check directly that

$$(10.11.9) \quad \begin{aligned} & \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} (A_{2, \mathcal{P}_2}(f))(x_1, x_2) dx_2 \right) dx_1 \\ &= \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} (A_{2, \mathcal{P}_2}(f))(x_1, x_2) dx_1 \right) dx_2. \end{aligned}$$

More precisely, one can verify that

$$(10.11.10) \quad \begin{aligned} & \int_{a_1}^{b_1} \left(\int_{t_{2,j-1}}^{t_{2,j}} (A_{2, \mathcal{P}_2}(f))(x_1, x_2) dx_2 \right) dx_1 \\ &= \int_{t_{2,j-1}}^{t_{2,j}} \left(\int_{a_1}^{b_1} (A_{2, \mathcal{P}_2}(f))(x_1, x_2) dx_1 \right) dx_2 \end{aligned}$$

for each $j = 1, \dots, k$, using the definition (10.10.4) of $A_{2, \mathcal{P}_2}(f)$. It is easy to see that (10.11.9) follows from (10.11.10), by summing over j .

10.11.4 Iterated integrals of f

We can use (10.11.9) to get that

$$(10.11.11) \quad \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x_1, x_2) dx_2 \right) dx_1 = \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x_1, x_2) dx_1 \right) dx_2,$$

as follows. The left side of (10.11.11) can be approximated by the left side of (10.11.9) when (10.10.11) is sufficiently small, because (10.10.6) is uniformly approximated by (10.10.7), as before. Similarly, the right side of (10.11.11) can be approximated by the right side of (10.11.9) when (10.10.11) is sufficiently small, because (10.11.3) is uniformly approximated by (10.11.5). This implies (10.11.11), because we can take \mathcal{P}_2 to be a partition of $[a_2, b_2]$ for which (10.10.11) is arbitrarily small.

10.12 A simpler approximation

Let us continue with the same notation and hypotheses as in the previous two sections. Put

$$(10.12.1) \quad a_{2, \mathcal{P}_2}(j) = \sup_{a_1 \leq x_1 \leq b_1} \sup_{t_{2,j-1} \leq x_2 \leq t_{2,j}} |f(x_1, x_2) - f(x_1, t_{2,j})|$$

for each $j = 1, \dots, k$, and

$$(10.12.2) \quad a_{2, \mathcal{P}_2} = \max_{1 \leq j \leq k} a_{2, \mathcal{P}_2}(j).$$

It is easy to see that a_{2, \mathcal{P}_2} is as small as we want when (10.10.11) is sufficiently small, because f is uniformly continuous on X .

10.12.1 Approximating some iterated integrals

Observe that

$$(10.12.3) \quad \begin{aligned} & \int_{a_1}^{b_1} \left(\int_{t_{2,j-1}}^{t_{2,j}} f(x_1, x_2) dx_2 \right) dx_1 - (t_{2,j} - t_{2,j-1}) \int_{a_1}^{b_1} f(x_1, t_{2,j}) dx_1 \\ &= \int_{a_1}^{b_1} \left(\int_{t_{2,j-1}}^{t_{2,j}} (f(x_1, x_2) - f(x_1, t_{2,j})) dx_2 \right) dx_1 \end{aligned}$$

for each $j = 1, \dots, k$. This implies that

$$(10.12.4) \quad \begin{aligned} & \left| \int_{a_1}^{b_1} \left(\int_{t_{2,j-1}}^{t_{2,j}} f(x_1, x_2) dx_2 \right) dx_1 - (t_{2,j} - t_{2,j-1}) \int_{a_1}^{b_1} f(x_1, t_{2,j}) dx_1 \right| \\ & \leq \int_{a_1}^{b_1} \left(\int_{t_{2,j-1}}^{t_{2,j}} |f(x_1, x_2) - f(x_1, t_{2,j})| dx_2 \right) dx_1 \\ & \leq (b_1 - a_1) (t_{2,j} - t_{2,j-1}) a_{2, \mathcal{P}_2}(j). \end{aligned}$$

Similarly,

$$(10.12.5) \quad \begin{aligned} & \left| \int_{t_{2,j-1}}^{t_{2,j}} \left(\int_{a_1}^{b_1} f(x_1, x_2) dx_1 \right) dx_2 - (t_{2,j} - t_{2,j-1}) \int_{a_1}^{b_1} f(x_1, t_{2,j}) dx_1 \right| \\ & \leq \int_{t_{2,j-1}}^{t_{2,j}} \left(\int_{a_1}^{b_1} |f(x_1, x_2) - f(x_1, t_{2,j})| dx_1 \right) dx_2 \\ & \leq (b_1 - a_1) (t_{2,j} - t_{2,j-1}) a_{2, \mathcal{P}_2}(j) \end{aligned}$$

for each $j = 1, \dots, k$.

10.12.2 Differences of some iterated integrals

Using (10.12.3) and (10.12.5), we get that

$$(10.12.6) \quad \begin{aligned} & \left| \int_{a_1}^{b_1} \left(\int_{t_{2,j-1}}^{t_{2,j}} f(x_1, x_2) dx_2 \right) dx_1 - \int_{t_{2,j-1}}^{t_{2,j}} \left(\int_{a_1}^{b_1} f(x_1, x_2) dx_1 \right) dx_2 \right| \\ & \leq 2(b_1 - a_1) (t_{2,j} - t_{2,j-1}) a_{2, \mathcal{P}_2}(j). \end{aligned}$$

It follows that

$$\begin{aligned}
 (10.12.7) \quad & \left| \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x_1, x_2) dx_2 \right) dx_1 - \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x_1, x_2) dx_1 \right) dx_2 \right| \\
 & \leq \sum_{j=1}^k 2(b_1 - a_1)(t_{2,j} - t_{2,j-1}) a_{2, \mathcal{P}_2}(j) \\
 & \leq 2(b_1 - a_1) \sum_{j=1}^k (t_{2,j} - t_{2,j-1}) a_{2, \mathcal{P}_2} \\
 & = 2(b_1 - a_1)(b_2 - a_2) a_{2, \mathcal{P}_2}.
 \end{aligned}$$

10.12.3 Equality of some iterated integrals

The right side of (10.12.7) is as small as we want when (10.10.11) is sufficiently small, as before. This is another way to obtain (10.11.11), by taking \mathcal{P}_2 to be a partition of $[a_2, b_2]$ for which (10.10.11) is arbitrarily small again.

10.13 Another approximation

Let us continue with the same notation and hypotheses as in the previous three sections. Also let $\mathcal{P}_1 = \{t_{1,l}\}_{l=0}^m$ be a partition of $[a_1, b_1]$, so that

$$(10.13.1) \quad a_1 = t_{1,0} < \cdots < t_{1,m} = b_1.$$

10.13.1 Iterated integrals as double sums

Of course,

$$\begin{aligned}
 (10.13.2) \quad & \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x_1, x_2) dx_2 \right) dx_1 \\
 & = \sum_{l=1}^m \sum_{j=1}^k \int_{t_{1,l-1}}^{t_{1,l}} \left(\int_{t_{2,j-1}}^{t_{2,j}} f(x_1, x_2) dx_2 \right) dx_1.
 \end{aligned}$$

Similarly,

$$\begin{aligned}
 (10.13.3) \quad & \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x_1, x_2) dx_1 \right) dx_2 \\
 & = \sum_{l=1}^m \sum_{j=1}^k \int_{t_{2,j-1}}^{t_{2,j}} \left(\int_{t_{1,l-1}}^{t_{1,l}} f(x_1, x_2) dx_1 \right) dx_2.
 \end{aligned}$$

10.13.2 Approximating the summands

If $1 \leq l \leq m$ and $1 \leq j \leq k$, then we can approximate

$$(10.13.4) \quad \int_{t_{1,l-1}}^{t_{1,l}} \left(\int_{t_{2,j-1}}^{t_{2,j}} f(x_1, x_2) dx_2 \right) dx_1$$

and

$$(10.13.5) \quad \int_{t_{2,j-1}}^{t_{2,j}} \left(\int_{t_{1,l-1}}^{t_{1,l}} f(x_1, x_2) dx_1 \right) dx_2$$

by

$$(10.13.6) \quad (t_{1,l} - t_{1,l-1}) (t_{2,j} - t_{2,j-1}) f(t_{1,l}, t_{2,j}).$$

If $t_{1,l} - t_{1,l-1}$ and $t_{2,j} - t_{2,j-1}$ are sufficiently small, then the errors in these approximations will be small compared to

$$(10.13.7) \quad (t_{1,l} - t_{1,l-1}) (t_{2,j} - t_{2,j-1}),$$

because of the uniform continuity of f on X . This means that the difference of (10.13.4) and (10.13.5) is small compared to (10.13.7) in this case.

10.13.3 Equality of the iterated integrals

Note that

$$(10.13.8) \quad \begin{aligned} & \sum_{l=1}^m \sum_{j=1}^k (t_{1,l} - t_{1,l-1}) (t_{2,j} - t_{2,j-1}) \\ &= \left(\sum_{l=1}^m (t_{1,l} - t_{1,l-1}) \right) \left(\sum_{j=1}^k (t_{2,j} - t_{2,j-1}) \right) \\ &= (b_1 - a_1) (b_2 - a_2). \end{aligned}$$

It follows that the difference of (10.13.2) and (10.13.3) is as small as we like when \mathcal{P}_1 and \mathcal{P}_2 are sufficiently fine partitions of $[a_1, b_1]$ and $[a_2, b_2]$, respectively.

This implies that the iterated integrals are equal, as in (10.11.11), because they do not depend on \mathcal{P}_1 or \mathcal{P}_2 . Basically, we are approximating the iterated integrals on the left sides of (10.13.2) and (10.13.3) by two-dimensional Riemann sums, and these two-dimensional Riemann sums are equal to each other.

10.14 Partitions of unity

Let (X, d_X) be a metric space. As in Section A.11, the support $\text{supp } f$ of a real or complex-valued function f on X is defined to be the closure in X of the set of $x \in X$ such that $f(x) \neq 0$.

It is often helpful to be able to find finite collections of continuous real-valued functions ψ_1, \dots, ψ_r on X with properties like the following. First,

$$(10.14.1) \quad 0 \leq \psi_j(x) \leq 1$$

for every $j = 1, \dots, r$. Second,

$$(10.14.2) \quad \sum_{j=1}^r \psi_j(x) = 1$$

for all x in a particular subset of X , which may be X itself. If (10.14.2) holds on a proper subset of X , then one may ask that

$$(10.14.3) \quad \sum_{j=1}^r \psi_j(x) \leq 1$$

for every $x \in X$. Of course, the first inequality in (10.14.1) together with (10.14.3) implies the second inequality in (10.14.1).

In addition, one may want to have restrictions on the supports of the ψ_j 's. One may ask that the supports of the ψ_j 's be contained in open sets in a particular family, for instance. A collection of functions like this is called a *partition of unity* on the set where (10.14.2) holds.

10.14.1 Some piecewise-linear functions

Let a, b be real numbers with $a < b$, and let $\mathcal{P} = \{t_j\}_{j=0}^k$ be a partition of $[a, b]$. If $0 \leq l \leq r$, then there is a unique nonnegative real-valued function τ_l on $[a, b]$ that is piecewise linear with breakpoints in \mathcal{P} such that

$$(10.14.4) \quad \begin{aligned} \tau_l(t_j) &= 1 & \text{when } j = l \\ &= 0 & \text{when } j \neq l. \end{aligned}$$

It is easy to see that

$$(10.14.5) \quad \sum_{l=0}^k \tau_l(x) = 1$$

for every $x \in [a, b]$, because the left side is a piecewise-linear function on $[a, b]$ with breakpoints in \mathcal{P} that is equal to 1 at t_j for each $j = 0, \dots, k$.

The support of τ_l , as a real-valued function on $[a, b]$, is given by

$$(10.14.6) \quad \begin{aligned} \text{supp } \tau_l &= [t_0, t_1] & \text{when } l = 0 \\ &= [t_{l-1}, t_{l+1}] & \text{when } 1 \leq l \leq k-1 \\ &= [t_{k-1}, t_k] & \text{when } l = k. \end{aligned}$$

10.14.2 Getting some partitions of unity

Let (X, d_X) be any metric space again, and let ϕ_1, \dots, ϕ_r be nonnegative continuous real-valued functions on X . There are a couple of common ways to get a partition of unity ψ_1, \dots, ψ_r from ϕ_1, \dots, ϕ_r , under suitable conditions.

In the first approach, we put

$$(10.14.7) \quad \Phi(x) = \sum_{j=1}^r \phi_j(x)$$

for each $x \in X$, and suppose that

$$(10.14.8) \quad \Phi(x) > 0 \quad \text{for every } x \in X.$$

In this case,

$$(10.14.9) \quad \psi_j(x) = \phi_j(x)/\Phi(x)$$

defines a continuous real-valued function on X for each $j = 1, \dots, r$, and (10.14.2) holds for every $x \in X$, by construction. We also have that

$$(10.14.10) \quad \text{supp } \psi_j = \text{supp } \phi_j$$

for every $j = 1, \dots, r$.

10.14.3 Another way to get a partition of unity

In the second approach, we suppose that

$$(10.14.11) \quad 0 \leq \phi_j(x) \leq 1$$

for every $j = 1, \dots, r$ and $x \in X$. Put $\psi_1 = \phi_1$, and

$$(10.14.12) \quad \psi_l = \left(\prod_{j=1}^{l-1} (1 - \phi_j) \right) \phi_l$$

for $l = 2, \dots, r$, as on p251 of [210]. Clearly (10.14.1) holds for every $j = 1, \dots, r$, and one can check that

$$(10.14.13) \quad \sum_{j=1}^l \psi_j = 1 - \prod_{j=1}^l (1 - \phi_j)$$

for every $l = 1, \dots, r$, by induction. This implies that (10.14.3) holds for every $x \in X$, and that (10.14.2) holds for every $x \in X$ such that

$$(10.14.14) \quad \phi_j(x) = 1 \text{ for some } j = 1, \dots, r.$$

Note that

$$(10.14.15) \quad \text{supp } \psi_j \subseteq \text{supp } \phi_j$$

for every $j = 1, \dots, r$.

10.15 Another approximation argument

Let (X_1, d_{X_1}) , (X_2, d_{X_2}) be metric spaces, and put $X = X_1 \times X_2$. Also let $d_{X,1}$, $d_{X,2}$ and $d_{X,\infty}$ be the corresponding metrics defined on X as in Subsection 10.1.1. As before, we shall refer to functions on X as being continuous if they are continuous with respect to any of these three metrics, and thus with respect to the other two metrics.

10.15.1 Using a partition of unity

Let $\psi_{2,1}, \dots, \psi_{2,r}$ be finitely many continuous nonnegative real-valued functions on X_2 that form a partition of unity on X_2 , so that

$$(10.15.1) \quad \sum_{j=1}^r \psi_{2,j}(x_2) = 1$$

for every $x_2 \in X_2$. If f is a real-valued function on X , then

$$(10.15.2) \quad f(x_1, x_2) = \sum_{j=1}^r f(x_1, x_2) \psi_{2,j}(x_2)$$

for every $x_1 \in X_1$ and $x_2 \in X_2$.

10.15.2 An approximation of f

Suppose that for each $j = 1, \dots, r$, $x_{2,j} \in X_2$ and

$$(10.15.3) \quad \psi_{2,j}(x_{2,j}) > 0.$$

Put

$$(10.15.4) \quad (A_2(f))(x_1, x_2) = \sum_{j=1}^r f(x_1, x_{2,j}) \psi_{2,j}(x_2)$$

for every $x_1 \in X_1$ and $x_2 \in X_2$. Thus

$$(10.15.5) \quad f(x_1, x_2) - (A_2(f))(x_1, x_2) = \sum_{j=1}^r (f(x_1, x_2) - f(x_1, x_{2,j})) \psi_{2,j}(x_2)$$

for every $x_1 \in X_1$ and $x_2 \in X_2$, by (10.15.2). It follows that

$$(10.15.6) \quad |f(x_1, x_2) - (A_2(f))(x_1, x_2)| \leq \sum_{j=1}^r |f(x_1, x_2) - f(x_1, x_{2,j})| \psi_{2,j}(x_2)$$

for every $x_1 \in X_1$ and $x_2 \in X_2$.

10.15.3 Some simple estimates

Put

$$(10.15.7) \quad a_{2,j}(x_1) = \sup\{|f(x_1, x_2) - f(x_1, x_{2,j})| : x_2 \in X_2, \psi_{2,j}(x_2) > 0\}$$

for each $j = 1, \dots, r$. The right side is allowed to be $+\infty$ here, although we shall typically be concerned with situations where it is finite. Observe that

$$(10.15.8) \quad |f(x_1, x_2) - f(x_1, x_{2,j})| \psi_{2,j}(x_2) \leq a_{2,j}(x_1) \psi_{2,j}(x_2)$$

for every $x_1 \in X_1$ and $x_2 \in X_2$, where the right side may be interpreted as being equal to 0 when $\psi_{2,j}(x_2) = 0$, even if (10.15.7) is $+\infty$. This implies that

$$(10.15.9) \quad |f(x_1, x_2) - (A_2(f))(x_1, x_2)| \leq \sum_{j=1}^r a_{2,j}(x_1) \psi_{2,j}(x_2)$$

for every $x_1 \in X_1$ and $x_2 \in X_2$, by (10.15.6).

Put

$$(10.15.10) \quad a_2(x_1) = \max_{1 \leq j \leq r} a_{2,j}(x_1)$$

for each $x_1 \in X_1$. Using (10.15.9), we get that

$$(10.15.11) \quad |f(x_1, x_2) - (A_2(f))(x_1, x_2)| \leq a_2(x_1)$$

for every $x_1 \in X_1$, because of (10.15.1).

10.15.4 Using these estimates

We may frequently be interested in situations where X_2 is compact, or at least totally bounded, and

$$(10.15.12) \quad \{x_2 \in X_2 : \psi_{2,j}(x_2) > 0\}$$

is a small subset of X_2 for each $j = 1, \dots, r$. This may mean that (10.15.12) is contained in a ball in X_2 centered at $x_{2,j}$ with small radius, or that (10.15.12) has small diameter with respect to d_{X_2} , which is nearly the same thing. If (10.15.12) is sufficiently small in this way, then we may be able to show that (10.15.7) is small, using suitable continuity properties of f .

Suppose that $\mathcal{E}_2 = \{f(x_1, \cdot) : x_1 \in X_1\}$ is uniformly equicontinuous as a collection of real-valued functions on X_2 . In this case, (10.15.10) is as small as we like when (10.15.12) is contained in a ball centered at $x_{2,j}$ of sufficiently small radius for each $j = 1, \dots, r$, or the diameter of (10.15.12) is sufficiently small for each $j = 1, \dots, r$, which is almost the same thing, as before.

If X_1 is compact, then \mathcal{E}_2 is equicontinuous at every $x_2 \in X_2$, as in Section 10.6. If X_2 is compact, then the equicontinuity of \mathcal{E}_2 at every point in X_2 implies that \mathcal{E}_2 is uniformly equicontinuous on X_2 , as in Subsection 5.2.1. If f is uniformly continuous on X with respect to $d_{X,1}$, $d_{X,2}$, or $d_{X,3}$, then it is easy to see that \mathcal{E}_2 is uniformly equicontinuous on X_2 . If X_1 and X_2 are compact, then X is compact with respect to $d_{X,1}$, $d_{X,2}$, and $d_{X,3}$, as in Subsection 10.4.1. In this case, continuity of f on X implies uniform continuity, as usual.

10.16 Some continuous functions

If c is any real number, then one can check that

$$(10.16.1) \quad \max(t, c) \text{ and } \min(t, c)$$

are Lipschitz functions of $t \in \mathbf{R}$, with constant 1. Of course, this uses the standard Euclidean metric on the real line.

Let (X, d_X) be a metric space, and let A be a nonempty subset of X . If $x \in X$, then the distance from x to A with respect to d_X is defined by

$$(10.16.2) \quad \text{dist}(x, A) = \inf\{d(x, a) : a \in A\}.$$

One can check that

$$(10.16.3) \quad \text{dist}(x, A) = 0$$

if and only if x is an element of the closure \overline{A} of A in X . One can also verify that

$$(10.16.4) \quad \text{dist}(x, \overline{A}) = \text{dist}(x, A)$$

for every $x \in X$.

It is easy to see that

$$(10.16.5) \quad \text{dist}(x, A) \leq d(x, w) + \text{dist}(w, A)$$

for every $x, w \in X$. Using this, one can check that $\text{dist}(x, A)$ is Lipschitz with constant 1, as a real-valued function on X .

10.16.1 Pairs of disjoint closed sets

Let A, B be nonempty disjoint closed subsets of X . Observe that

$$(10.16.6) \quad \text{dist}(x, A) + \text{dist}(x, B) > 0$$

for every $x \in X$. It follows that

$$(10.16.7) \quad \frac{\text{dist}(x, A)}{\text{dist}(x, A) + \text{dist}(x, B)}$$

defines a continuous real-valued function on X . This function is equal to 0 exactly on A , and it is equal to 1 exactly on B . It also takes values in $[0, 1]$ on X , by construction.

10.16.2 Some uniformly continuous functions

Let $r > 0$ be given, and note that

$$(10.16.8) \quad \min(\text{dist}(x, A), r)$$

is Lipschitz with constant 1, as a real-valued function of $x \in X$. This implies that

$$(10.16.9) \quad (1/r) \min(\text{dist}(x, A), r)$$

is Lipschitz with constant $1/r$ on X .

Suppose for the moment that

$$(10.16.10) \quad d(a, b) \geq r$$

for every $a \in A$ and $b \in B$. Equivalently, this means that

$$(10.16.11) \quad \text{dist}(b, A) \geq r$$

for every $b \in B$. This implies that (10.16.8) is equal to r when $x \in B$, so that (10.16.9) is equal to 1.

Conversely, suppose that there is a uniformly continuous real-valued function f on X such that

$$(10.16.12) \quad f(a) = 0 \text{ for every } a \in A, \text{ and } f(b) = 1 \text{ for every } b \in B.$$

Under these conditions, it is easy to see that there is an $r > 0$ such that (10.16.10) holds for every $a \in A$ and $b \in B$.

10.16.3 Some functions with small support

Let $a_0 \in X$ and $r_0 > 0$ be given, and note that

$$(10.16.13) \quad d(x, a_0)$$

is Lipschitz with constant 1 as a real-valued function of $x \in X$. This is the same as (10.16.2), with $A = \{a_0\}$. It follows that

$$(10.16.14) \quad r_0 - d(x, a_0)$$

is Lipschitz with constant 1 on X , so that

$$(10.16.15) \quad \max(r_0 - d(x, a_0), 0)$$

is Lipschitz with constant 1 on X as well. Of course, (10.16.15) is equal to 0 when

$$(10.16.16) \quad d(x, a_0) \geq r_0,$$

and otherwise (10.16.15) is strictly positive.

10.16.4 Some functions equal to 1 near a_0

Let r_1 be a nonnegative real number with

$$(10.16.17) \quad r_1 < r_0,$$

and consider

$$(10.16.18) \quad \min(\max(r_0 - d(x, a_0), 0), r_0 - r_1).$$

This is another real-valued Lipschitz function on X with constant 1, so that

$$(10.16.19) \quad (1/(r_0 - r_1)) \min(\max(r_0 - d(x, a_0), 0), r_0 - r_1)$$

is Lipschitz with constant $1/(r_0 - r_1)$ on X . This function is equal to 0 when (10.16.16) holds, and it is equal to 1 when

$$(10.16.20) \quad d(x, a_0) \leq r_1.$$

If

$$(10.16.21) \quad r_1 < d(x, a_0) < r_0,$$

then (10.16.19) is in $(0, 1)$.

10.17 Products and reciprocals

Let (X, d_X) be a metric space, and let f, g be real-valued functions on X . If f and g are uniformly continuous on X , then it is easy to see that $f + g$ is uniformly continuous on X too. If f and g are Lipschitz functions on X , then $f + g$ is a Lipschitz function on X as well, as in Subsection 7.7.3.

Suppose for the moment that f and g are bounded on X . If f and g are also uniformly continuous on X , then one can check that their product fg is uniformly continuous on X . Similarly, if f and g are Lipschitz functions on X , then one can verify that fg is Lipschitz on X .

If t_1, t_2 are nonzero real numbers, then

$$(10.17.1) \quad 1/t_1 - 1/t_2 = (t_2 - t_1)/(t_1 t_2).$$

This implies that

$$(10.17.2) \quad |1/t_1 - 1/t_2| = |t_1 - t_2|/(|t_1| |t_2|).$$

If r is a positive real number, and $|t_1|, |t_2| \geq r$, then we get that

$$(10.17.3) \quad |1/t_1 - 1/t_2| \leq r^{-2} |t_1 - t_2|.$$

This means that $t \mapsto 1/t$ is Lipschitz with constant r^{-2} as a real-valued function on

$$(10.17.4) \quad \{t \in \mathbf{R} : |t| \geq r\}.$$

Of course, this uses the standard Euclidean metric on \mathbf{R} , and its restriction to (10.17.4).

Suppose now that

$$(10.17.5) \quad |f(x)| \geq r$$

for every $x \in X$. If f is uniformly continuous on X , then it is easy to see that

$$(10.17.6) \quad 1/f \text{ is uniformly continuous on } X.$$

Similarly, if f is Lipschitz on X , then

$$(10.17.7) \quad 1/f \text{ is Lipschitz on } X.$$

These statements can be verified directly, or by considering $1/f$ as the composition of f with $t \mapsto 1/t$ on $(10.17.4)$. There are analogous statements for complex numbers, and complex-valued functions on X .

10.18 Graphs of mappings

Let X_1, X_2 be nonempty sets, and let $X = X_1 \times X_2$ be their Cartesian product. If f is a mapping from X_1 into X_2 , then the *graph* of f is the subset of X given by

$$(10.18.1) \quad \{(x_1, f(x_1)) : x_1 \in X_1\},$$

as usual. Put

$$(10.18.2) \quad F(x_1) = (x_1, f(x_1))$$

for every $x_1 \in X_1$, which defines F as a mapping from X_1 into X . Note that (10.18.1) is the same as the image of X_1 under F .

Let p_1, p_2 be the usual coordinate projections from X onto X_1, X_2 , respectively, as in Section 10.5. Thus $p_j(x) = x_j$ for each $x = (x_1, x_2) \in X$ and $j = 1, 2$. Clearly

$$(10.18.3) \quad f = p_2 \circ F$$

on X_1 , and $p_1 \circ F$ is the identity mapping on X_1 .

Suppose now that (X_1, d_{X_1}) and (X_2, d_{X_2}) are metric spaces, so that we can define the metrics $d_{X,1}, d_{X,2}$, and $d_{X,\infty}$ on X as in Subsection 10.1.1. Remember that these metrics determine the same collections of open sets, closed sets, and compact sets in X , and that convergence of sequences in X with respect to these metrics are equivalent, as before.

If

$$(10.18.4) \quad f \text{ is continuous as a mapping from } X_1 \text{ into } X_2,$$

then one can check that

$$(10.18.5) \quad F \text{ is continuous as a mapping from } X_1 \text{ into } X,$$

with respect to each of $d_{X,1}, d_{X,2}$, and $d_{X,\infty}$ on X . Conversely, if (10.18.5) holds, with respect to any of the metrics $d_{X,1}, d_{X,2}$, or $d_{X,\infty}$ on X , then (10.18.4) holds. This can be seen using (10.18.3), and the fact that p_2 is a continuous mapping from X onto X_2 with respect to each of $d_{X,1}, d_{X,2}$, and $d_{X,\infty}$ on X , as in Section 10.5.

10.18.1 Closed graphs in X

One can check that a subset E of a metric space M is a closed set if and only if for every sequence $\{w_j\}_{j=1}^{\infty}$ of elements of E that converges to an element w of M , we have that

$$(10.18.6) \quad w \in E.$$

More precisely, if $\{w_j\}_{j=1}^{\infty}$ is a sequence of elements of any set $E \subseteq M$ that converges to an element w of M , and if $w_j \neq w$ for each j , then w is a limit point of E . If $w \in M$ is a limit point of E , then one can find a sequence $\{w_j\}_{j=1}^{\infty}$ of elements of E that converges to w , with $w_j \neq w$ for each j .

It is well known that if (10.18.4) holds, then

$$(10.18.7) \quad \text{the graph of } f \text{ is a closed set in } X,$$

with respect to each of the metrics $d_{X,1}$, $d_{X,2}$, $d_{X,\infty}$. As in the preceding paragraph, (10.18.7) holds if and only if for every sequence of elements of the graph of f that converges to an element of X , the limit of the sequence is in the graph of f . In this case, this means that (10.18.7) holds if and only if for every sequence $\{x_{1,j}\}_{j=1}^{\infty}$ of elements of X_1 such that

$$(10.18.8) \quad \{(x_{1,j}, f(x_{1,j}))\}_{j=1}^{\infty}$$

converges to an element (x_1, x_2) of X , we have that

$$(10.18.9) \quad f(x_1) = x_2.$$

As in Subsection 10.3.1, the convergence of (10.18.8) to $(x_1, x_2) \in X$ with respect to any of the metrics $d_{X,1}$, $d_{X,2}$, $d_{X,\infty}$ is equivalent to the convergence of $\{x_{1,j}\}_{j=1}^{\infty}$ to x_1 in X_1 and the convergence of $\{f(x_{1,j})\}_{j=1}^{\infty}$ to x_2 in X_2 . If f is continuous at x_1 , then the convergence of $\{x_{1,j}\}_{j=1}^{\infty}$ to x_1 in X_1 implies that $\{f(x_{1,j})\}_{j=1}^{\infty}$ converges to $f(x_1)$ in X_2 . This implies that (10.18.9) holds when $\{f(x_{1,j})\}_{j=1}^{\infty}$ converges to x_2 in X_2 .

10.18.2 Compact graphs in X

If (10.18.4) holds, and if

$$(10.18.10) \quad X_1 \text{ is compact,}$$

then

$$(10.18.11) \quad \text{the graph of } f \text{ is a compact subset of } X,$$

with respect to each of $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$. Remember that the graph of f is the same as $F(X_1)$, and that F is continuous as a mapping from X_1 into X in this case. This implies that $F(X_1)$ is a compact subset of X when X_1 is compact.

Conversely, let f be any mapping from X_1 into X_2 , and suppose that (10.18.11) holds, with respect to $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$. It is easy to see that (10.18.10) holds in this case, because p_1 maps the graph of f onto X_1 . This also

use the fact that p_1 is continuous as a mapping from X onto X_1 , as in Section 10.5.

It is well known that (10.18.4) holds under these conditions as well. To see this, let $x_1 \in X_1$ be given, and suppose for the sake of a contradiction that f is not continuous at x_1 . This means that there is an $\epsilon > 0$ such that for every $\delta > 0$ there is a point $w_1 \in X_1$ with

$$(10.18.12) \quad d_{X_1}(x_1, w_1) < \delta$$

and

$$(10.18.13) \quad d_{X_2}(f(x_1), f(w_1)) \geq \epsilon.$$

One can use this to get a sequence $\{x_{1,j}\}_{j=1}^{\infty}$ of elements of X_1 that converges to x_1 , with

$$(10.18.14) \quad d_{X_2}(f(x_1), f(x_{1,j})) \geq \epsilon$$

for every j .

If (10.18.11) holds, then the graph of f is sequentially compact in X . Using this, we get that there is a subsequence $\{x_{1,j_l}\}_{l=1}^{\infty}$ of $\{x_{1,j}\}_{j=1}^{\infty}$ such that

$$(10.18.15) \quad \{(x_{1,j_l}, f(x_{1,j_l}))\}_{l=1}^{\infty}$$

converges to an element $(y_1, f(y_1))$ of the graph of f in X . This means that

$$(10.18.16) \quad \{x_{1,j_l}\}_{l=1}^{\infty} \text{ converges to } y_1 \text{ in } X_1,$$

and that

$$(10.18.17) \quad \{f(x_{1,j_l})\}_{l=1}^{\infty} \text{ converges to } f(y_1) \text{ in } X_2,$$

as in Subsection 10.3.1. Note that $\{x_{1,j_l}\}_{l=1}^{\infty}$ converges to x_1 in X_1 , because $\{x_{1,j}\}_{j=1}^{\infty}$ converges to x_1 , by construction. Thus

$$(10.18.18) \quad x_1 = y_1.$$

It follows that

$$(10.18.19) \quad \{f(x_{1,j_l})\}_{l=1}^{\infty} \text{ converges to } f(x_1) \text{ in } X_2.$$

This contradicts (10.18.14), as desired.

10.19 Semicontinuity

Let us continue with the same notation and hypotheses as in the previous section, except that now we take $X_2 = \mathbf{R}$, with the standard Euclidean metric. If f is a real-valued function on X_1 , then

$$(10.19.1) \quad \{(x_1, x_2) \in X_1 \times \mathbf{R} : f(x_1) > x_2\}$$

and

$$(10.19.2) \quad \{(x_1, x_2) \in X_1 \times \mathbf{R} : f(x_1) \leq x_2\}$$

are complementary subsets of $X = X_1 \times \mathbf{R}$. In particular, (10.19.1) is an open set in $X = X_1 \times \mathbf{R}$ if and only if (10.19.2) is a closed set in X , with respect to any, and thus each, of the metrics $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$.

Similarly,

$$(10.19.3) \quad \{(x_1, x_2) \in X_1 \times \mathbf{R} : f(x_1) \geq x_2\}$$

and

$$(10.19.4) \quad \{(x_1, x_2) \in X_1 \times \mathbf{R} : f(x_1) < x_2\}$$

are complementary subsets of X . It follows that (10.19.3) is a closed set in X if and only if (10.19.4) is an open set, as before. Note that the graph of f is the same as the intersection of (10.19.2) and (10.19.3).

If f is continuous on X , then one can show that (10.19.2) and (10.19.3) are closed sets in X , using the same type of argument as in Subsection 10.18.1. However, there are more precise statements, using notions of *semicontinuity*.

10.19.1 Upper and lower semicontinuity

We say that f is *upper semicontinuous* at a point $x_1 \in X_1$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$(10.19.5) \quad f(w_1) < f(x_1) + \epsilon$$

for every $w_1 \in X_1$ with $d_{X_1}(x_1, w_1) < \delta$. Similarly, we say that f is *lower semicontinuous* at x_1 if for every $\epsilon > 0$ there is a $\delta' > 0$ such that

$$(10.19.6) \quad f(w_1) > f(x_1) - \epsilon$$

for every $w_1 \in X_1$ with $d_{X_1}(x_1, w_1) < \delta'$.

Observe that f is continuous at x_1 if and only if f is both upper and lower semicontinuous at x_1 . It is easy to see that f is upper semicontinuous at x_1 if and only if $-f$ is lower semicontinuous at x_1 .

Let us say that f is *upper semicontinuous* on X_1 if f is upper semicontinuous at every $x_1 \in X_1$. Similarly, we say that f is *lower semicontinuous* on X_1 if f is lower semicontinuous at every $x_1 \in X_1$.

One can check that f is upper semicontinuous on X_1 if and only if for every real number b ,

$$(10.19.7) \quad \{x_1 \in X_1 : f(x_1) < b\}$$

is an open subset of X_1 . Similarly, f is lower semicontinuous on X_1 if and only if for every $a \in \mathbf{R}$,

$$(10.19.8) \quad \{x_1 \in X_1 : f(x_1) > a\}$$

is an open set in X_1 .

One can also verify that f is upper semicontinuous on X_1 if and only if (10.19.4) is an open set in X . Similarly, f is lower semicontinuous on X_1 if and only if (10.19.1) is an open set in X .

Let K be a nonempty compact subset of X_1 . If f is upper semi-continuous on X_1 , then it is not too difficult to show that f attains its maximum on K .

Similarly, if f is lower semicontinuous on X_1 , then f attains its minimum on K . Of course, this is another version of the *extreme value theorem*.

Let $\{x_{1,j}\}_{j=1}^{\infty}$ be a sequence of elements of X_1 that converges to $x_1 \in X_1$. If f is upper semi-continuous at x_1 , then one can check that

$$(10.19.9) \quad \limsup_{j \rightarrow \infty} f(x_j) \leq f(x_1).$$

Similarly, if f is lower semicontinuous at x_1 , then

$$(10.19.10) \quad \liminf_{j \rightarrow \infty} f(x_j) \geq f(x_1).$$

One can verify that these properties characterize upper and lower semicontinuity of f at x_1 too.

10.20 Homeomorphisms between metric spaces

Let (X, d_X) and (Y, d_Y) be metric spaces. A one-to-one mapping f from X onto Y is said to be a *homeomorphism* if f is continuous, and the corresponding inverse mapping f^{-1} from Y onto X is continuous. Of course, f^{-1} is a homeomorphism from Y onto X in this case.

Let (Z, d_Z) be another metric space. If f is a homeomorphism from X onto Y , and g is a homeomorphism from Y onto Z , then their composition $g \circ f$ is a homeomorphism from X onto Z .

10.20.1 A criterion using compactness

Let f be a one-to-one continuous mapping from X onto Y . If X is compact, then it is well known that the inverse mapping f^{-1} is continuous, so that

$$(10.20.1) \quad f \text{ is a homeomorphism.}$$

One can show this using sequences, or by looking at closed sets.

To show this using sequences, let $\{y_j\}_{j=1}^{\infty}$ be a sequence of elements of Y that converges to a point $y \in Y$, and let us check that

$$(10.20.2) \quad \{f^{-1}(y_j)\}_{j=1}^{\infty} \text{ converges to } f^{-1}(y) \text{ in } X.$$

Put $x_j = f^{-1}(y_j)$ for each j , and $x = f^{-1}(y)$, for convenience. Suppose for the sake of a contradiction that $\{x_j\}_{j=1}^{\infty}$ does not converge to x in X . This implies that there is an $\epsilon > 0$ such that

$$(10.20.3) \quad d_X(x, x_j) \geq \epsilon$$

for infinitely many $j \geq 1$. Equivalently, this means that there is a subsequence $\{x_{j_l}\}_{l=1}^{\infty}$ of $\{x_j\}_{j=1}^{\infty}$ such that

$$(10.20.4) \quad d_X(x, x_{j_l}) \geq \epsilon$$

for every $l \geq 1$.

If X is compact, and thus sequentially compact, then there is a subsequence $\{x_{j_{l_n}}\}_{n=1}^\infty$ of $\{x_{j_l}\}_{l=1}^\infty$ that converges to an element w of X . It follows that

$$(10.20.5) \quad \{f(x_{j_{l_n}})\}_{n=1}^\infty \text{ converges to } f(w) \text{ in } Y,$$

because f is continuous at w , by hypothesis. However,

$$(10.20.6) \quad \{f(x_{j_{l_n}})\}_{n=1}^\infty = \{y_{j_{l_n}}\}_{n=1}^\infty$$

is a subsequence of $\{y_j\}_{j=1}^\infty$, which converges to $y = f(x)$. This implies that

$$(10.20.7) \quad f(w) = f(x),$$

so that $w = x$, because f is one-to-one. This means that $\{x_{j_{l_n}}\}_{n=1}^\infty$ converges to x in X , contradicting (10.20.4), as desired.

10.20.2 An argument using closed sets

Alternatively, let E be a closed set in X . In order to check that f^{-1} is continuous as a mapping from Y into X , we would like to verify that

$$(10.20.8) \quad (f^{-1})^{-1}(E) = \{y \in Y : f^{-1}(y) \in E\}$$

is a closed set in Y . It is easy to see that

$$(10.20.9) \quad (f^{-1})^{-1}(E) = f(E).$$

Note that

$$(10.20.10) \quad E \text{ is compact in } X,$$

because X is compact, and E is a closed set. This implies that

$$(10.20.11) \quad f(E) \text{ is a compact set in } Y,$$

because f is continuous. It follows that

$$(10.20.12) \quad f(E) \text{ is a closed set in } Y.$$

This means that f^{-1} is continuous, by a standard characterization of continuous mappings between metric spaces.

One might notice that the first argument is very similar to one in Subsection 10.18.2. This will be discussed further in the next section.

10.21 Graphs and homeomorphisms

Let (X_1, d_{X_1}) , (X_2, d_{X_2}) be nonempty metric spaces, and put $X = X_1 \times X_2$. As usual, we can define metrics $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ on X as in Subsection 10.1.1.

Let f be a mapping from X_1 into X_2 , and let F be the mapping from X_1 into X defined by putting $F(x_1) = (x_1, f(x_1))$ for every $x_1 \in X_1$, as in Section 10.18.

Let

$$(10.21.1) \quad Y = F(X_1)$$

be the graph of f in X , as before. We may consider Y as a metric space, using the restriction of $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$ to Y . Note that F is a one-to-one mapping from X_1 onto Y , by construction.

Let p_1, p_2 be the usual coordinate projections from X onto X_1, X_2 , respectively, as in Section 10.5. The restriction of p_1 to Y is the same as the inverse mapping F^{-1} of F on Y . It follows that

$$(10.21.2) \quad F^{-1} \text{ is continuous on } Y,$$

with respect to the restriction of $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$ to Y , as in Section 10.5.

If f is continuous as a mapping from X_1 into X_2 , then F is continuous as a mapping from X_1 into X with respect to each of $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$, as in Section 10.18. In this case, we get that

$$(10.21.3) \quad F \text{ is a homeomorphism from } X_1 \text{ onto } Y,$$

with respect to the restriction of any of $d_{X,1}$, $d_{X,2}$, and $d_{X,\infty}$ to Y .

If Y is compact with respect to any of $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$, then one can get that

$$(10.21.4) \quad F \text{ is continuous,}$$

because F^{-1} is continuous on Y , as in the previous section. Of course, this implies that

$$(10.21.5) \quad f = p_2 \circ F \text{ is continuous on } X_1,$$

because p_2 is continuous on X , as in Section 10.5. This is another way to look at how the compactness of Y implies the continuity of f , as in Subsection 10.18.2.

10.21.1 The case where $X_2 = \mathbf{R}^m$

Let m be a positive integer, and let us now take

$$(10.21.6) \quad X_2 = \mathbf{R}^m,$$

equipped with the standard Euclidean metric, or the metric associated to a norm. Let f be a mapping from X_1 into \mathbf{R}^m , and let Φ be the mapping from $X = X_1 \times \mathbf{R}^m$ into itself defined by

$$(10.21.7) \quad \Phi(x) = (x_1, x_2 + f(x_1))$$

for every $x = (x_1, x_2) \in X$. It is easy to see that Φ is a one-to-one mapping from X onto itself. More precisely, the inverse mapping is given by

$$(10.21.8) \quad \Phi^{-1}(x) = (x_1, x_2 - f(x_1))$$

for every $x \in X$.

If f is continuous on X_1 , then one can check that Φ and Φ^{-1} are continuous on X , with respect to any of the metrics $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$. This means that Φ is a homeomorphism from X onto itself, with respect to any of these three metrics. It is easy to see that the continuity of f is necessary for the continuity of Φ or Φ^{-1} .

If $x_1 \in X_1$ and $x = (x_1, 0)$, then $\Phi(x) = F(x_1)$. In particular,

$$(10.21.9) \quad \Phi(X_1 \times \{0\}) = F(X_1).$$

Equivalently,

$$(10.21.10) \quad F(X_1) = (\Phi^{-1})^{-1}(X_1 \times \{0\}).$$

It is easy to see that $X_1 \times \{0\}$ is a closed set in X , with respect to any of $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$. If f is continuous, then Φ^{-1} is continuous, and (10.21.10) gives another way to see that $F(X_1)$ is a closed set in X , as in Subsection 10.18.1.

Suppose now that $m = 1$, so that f is a real-valued function on X_1 . Observe that

$$(10.21.11) \quad \Phi(X_1 \times (0, +\infty)) = \{(x_1, x_2) \in X_1 \times \mathbf{R} : x_2 > f(x_1)\}$$

and

$$(10.21.12) \quad \Phi(X_1 \times (-\infty, 0)) = \{(x_1, x_2) \in X_1 \times \mathbf{R} : x_2 < f(x_1)\}.$$

It is easy to see that $X_1 \times (0, +\infty)$ and $X_1 \times (-\infty, 0)$ are open sets in X , with respect to any of $d_{X,1}$, $d_{X,2}$, or $d_{X,\infty}$. If f is continuous, then one can use the continuity of Φ^{-1} to get another way to see that (10.21.11) and (10.21.12) are open sets in X , as in Section 10.19.

Chapter 11

Summable functions

Sums of real and complex-valued functions on arbitrary nonempty sets are considered in this chapter, extending sums of absolutely convergent series of real and complex numbers. The reader may choose to skip this chapter, at least initially.

The types of sums discussed in this chapter may also be considered as Lebesgue integrals with respect to counting measure. Some of the arguments used here are analogous to ones commonly employed in the study of Lebesgue integrals, although simpler.

11.1 Extended real numbers

As usual, the set of *extended real numbers* consists of the real numbers together with two additional elements, denoted $+\infty$ and $-\infty$. The standard ordering is extended to the set of extended real numbers by putting

$$(11.1.1) \quad -\infty < x < +\infty$$

for every $x \in \mathbf{R}$. Normally when we consider extended real numbers here, we shall only be concerned with nonnegative extended real numbers.

11.1.1 Addition and multiplication

In some situations, addition and multiplication of extended real numbers can be defined in a natural way. In particular, we put

$$(11.1.2) \quad x + \infty = \infty + x = \infty + \infty = +\infty$$

for every $x \in \mathbf{R}$. Similarly, we put

$$(11.1.3) \quad x \cdot \infty = \infty \cdot x = \infty \cdot \infty = \infty$$

for every positive real number x . Although $0 \cdot \infty$ is not necessarily defined, it will normally correspond to 0 here.

11.1.2 Upper and lower bounds

The notions of upper and lower bounds, supremum, and infimum can be defined for sets of extended real numbers, in the same way as for sets of real numbers. In particular, if A is a nonempty set of real numbers that does not have a finite upper bound in \mathbf{R} , then the supremum of A can be defined as an extended real number to be $+\infty$. Of course, if $+\infty$ is an element of A , then the supremum of A is equal to $+\infty$ automatically. If $+\infty$ is the only element of A , then the infimum of A is equal to $+\infty$.

Let t be a positive real number, and let E be a nonempty set of extended real numbers. Put

$$(11.1.4) \quad tE = \{tx : x \in E\},$$

which is another nonempty set of extended real numbers. One can check that

$$(11.1.5) \quad \sup(tE) = t(\sup E).$$

11.1.3 Sequences tending to $+\infty$

Let $\{x_j\}_{j=1}^{\infty}$ be a sequence of real numbers. If $\{x_j\}_{j=1}^{\infty}$ converges to a real number x with respect to the standard metric on \mathbf{R} , then we may express this by $x_j \rightarrow x$ as $j \rightarrow \infty$. As usual, we say that $\{x_j\}_{j=1}^{\infty}$ tends to $+\infty$ as $j \rightarrow \infty$, or $x_j \rightarrow +\infty$ as $j \rightarrow \infty$, if for every positive real number R there is a positive integer L such that

$$(11.1.6) \quad x_j > R$$

for each $j \geq L$.

11.1.4 Monotonically increasing sequences

Let $\{x_j\}_{j=1}^{\infty}$ be a monotonically increasing sequence of real numbers, so that

$$(11.1.7) \quad x_j \leq x_{j+1}$$

for every $j \geq 1$. Put

$$(11.1.8) \quad x = \sup\{x_j : j \in \mathbf{Z}_+\},$$

which is a real number when the set of x_j 's, $j \in \mathbf{Z}_+$, has a finite upper bound in \mathbf{R} , and otherwise is equal to $+\infty$. It is well known that

$$(11.1.9) \quad x_j \rightarrow x \quad \text{as } j \rightarrow \infty$$

under these conditions.

Now let $\{x_j\}_{j=1}^{\infty}$ be a sequence of real numbers such that $x_j \rightarrow +\infty$ as $j \rightarrow \infty$. If $\{y_j\}_{j=1}^{\infty}$ is a sequence of real numbers with a finite lower bound in \mathbf{R} , then it is easy to see that

$$(11.1.10) \quad x_j + y_j \rightarrow +\infty \quad \text{as } j \rightarrow \infty.$$

In particular, this holds when $y_j \rightarrow y$ as $j \rightarrow \infty$, where $y \in \mathbf{R}$ or $y = +\infty$.

11.2 Nonnegative sums

Let X be a nonempty set, and let f be a nonnegative real-valued function on X . If A is a nonempty finite subset of X , then the sum

$$(11.2.1) \quad \sum_{x \in A} f(x)$$

can be defined as a nonnegative real number in the usual way.

11.2.1 Summing nonnegative real-valued functions

Put

$$(11.2.2) \quad \sum_{x \in X} f(x) = \sup \left\{ \sum_{x \in A} f(x) : A \text{ is a nonempty finite subset of } X \right\},$$

where the supremum on the right side is defined as a nonnegative extended real number, as in Subsection 11.1.2. Of course, if X has only finitely many elements, then the supremum is attained with $A = X$. Similarly, if f has finite support in X , then the supremum is attained with any nonempty finite subset A of X such that A contains the support of f .

11.2.2 The case where $X = \mathbf{Z}_+$

Suppose for the moment that X is the set \mathbf{Z}_+ of positive integers, and let f be a nonnegative real-valued function defined on \mathbf{Z}_+ . In this case, one may put

$$(11.2.3) \quad \sum_{j=1}^{\infty} f(j) = \sup \left\{ \sum_{j=1}^n f(j) : n \in \mathbf{Z}_+ \right\},$$

where the supremum on the right side is defined as a nonnegative extended real number again.

Of course, the sequence of partial sums $\sum_{j=1}^n f(j)$ increases monotonically in this case. The sequence of partial sums tends to its supremum as $n \rightarrow \infty$, as in Subsection 11.1.4.

If the sequence of partial sums has a finite upper bound in \mathbf{R} , then the sequence of partial sums converges in \mathbf{R} in the usual sense. This means that the infinite series on the left side of (11.2.3) converges in the usual sense, with sum equal to the right side of (11.2.3). If the sequence of partial sums does not have a finite upper bound, so that the right side of (11.2.3) is $+\infty$, then one may interpret the sum on the left side of (11.2.3) as being $+\infty$ as well.

11.2.3 Equivalence of the two definitions in this case

If f is any nonnegative real-valued function on \mathbf{Z}_+ , then

$$(11.2.4) \quad \sum_{j=1}^{\infty} f(j) = \sum_{j \in \mathbf{Z}_+} f(j),$$

where these sums are as defined in (11.2.2) and (11.2.3). More precisely, one can check that

$$(11.2.5) \quad \sum_{j=1}^{\infty} f(j) \leq \sum_{j \in \mathbf{Z}_+} f(j),$$

because each of the partial sums on the right side of (11.2.3) may be considered as one of the finite sums on the right side of (11.2.2). To get the other inequality, one can use the fact that every finite subset A of \mathbf{Z}_+ is contained in $\{1, \dots, n\}$ for some $n \in \mathbf{Z}_+$.

11.2.4 A sometimes convenient extension

It is sometimes convenient to consider a nonnegative extended real-valued function f on a nonempty set X . If A is a nonempty finite subset of X and $f(x) = +\infty$ for some $x \in A$, then the corresponding sum (11.2.1) is equal to $+\infty$. If $f(x) = +\infty$ for some $x \in X$, then the right side of (11.2.2) is equal to $+\infty$. Similarly, if $X = \mathbf{Z}_+$, then the partial sum $\sum_{j=1}^n f(j)$ is $+\infty$ when $f(j) = +\infty$ for some $j \leq n$, so that the right side of (11.2.3) is $+\infty$ when $f(j) = +\infty$ for some j . Thus (11.2.4) also holds in this situation.

11.3 Compositions and subsets

Let X and Y be nonempty sets, and let ϕ be a one-to-one mapping from X onto Y . Also let f be a nonnegative extended real-valued function on Y , so that $f(\phi(x))$ defines a nonnegative extended real-valued function on X . If A is a nonempty finite subset of X , then $\phi(A)$ is a nonempty finite subset of Y , and

$$(11.3.1) \quad \sum_{x \in A} f(\phi(x)) = \sum_{y \in \phi(A)} f(y).$$

If B is any nonempty finite subset of Y , then $A = \phi^{-1}(B)$ is a nonempty finite subset of X , and $B = \phi(A)$. It follows that

$$(11.3.2) \quad \sum_{x \in X} f(\phi(x)) = \sum_{y \in Y} f(y),$$

because both sums are defined by taking the supremum of the corresponding finite subsums in (11.3.1).

11.3.1 Sums over subsets

Let f be a nonnegative extended real-valued function on a nonempty set X . If E is a nonempty subset of X , then the sum

$$(11.3.3) \quad \sum_{x \in E} f(x)$$

can be defined as a nonnegative extended real number in the same way as before, as the supremum of the corresponding collection of finite subsums. This is the same as applying the definition in the previous section to the restriction of f to E .

If E_1, E_2 are nonempty subsets of X and $E_1 \subseteq E_2$, then

$$(11.3.4) \quad \sum_{x \in E_1} f(x) \leq \sum_{x \in E_2} f(x).$$

This uses the fact that every finite subsum of the sum on the left is also a finite subsum of the sum on the right. If we also have that $f(x) = 0$ for every $x \in E_2 \setminus E_1$, then it follows that

$$(11.3.5) \quad \sum_{x \in E_1} f(x) = \sum_{x \in E_2} f(x).$$

This is because finite subsums of the sum on the right side can be reduced to finite subsums of the sum on the left, except possibly for sums over nonempty finite subsets of $E_2 \setminus E_1$, which are equal to 0 in this case.

11.3.2 Listing terms with a sequence

Let $\{x_j\}_{j=1}^{\infty}$ be a sequence of distinct elements of X , and let

$$(11.3.6) \quad E = \{x_j : j \in \mathbf{Z}_+\}.$$

If f is a nonnegative extended real-valued function on X again, then $f(x_j)$ may be considered as a nonnegative extended real-valued function on \mathbf{Z}_+ . Thus

$$(11.3.7) \quad \sum_{j=1}^{\infty} f(x_j) = \sum_{j \in \mathbf{Z}_+} f(x_j),$$

as in (11.2.4). We also have that

$$(11.3.8) \quad \sum_{j \in \mathbf{Z}_+} f(x_j) = \sum_{x \in E} f(x),$$

as in (11.3.2). More precisely, this uses the fact that $j \mapsto x_j$ is a one-to-one mapping from \mathbf{Z}_+ onto E .

11.3.3 Monotonicity of the sum

Let f, g be nonnegative extended real-valued functions on X , and suppose that

$$(11.3.9) \quad f(x) \leq g(x)$$

for every $x \in X$. If A is a nonempty finite subset of X , then

$$(11.3.10) \quad \sum_{x \in A} f(x) \leq \sum_{x \in A} g(x).$$

Using this, it is easy to see that

$$(11.3.11) \quad \sum_{x \in X} f(x) \leq \sum_{x \in X} g(x).$$

11.4 Nonnegative summable functions

Let f be a nonnegative real-valued function on a nonempty set X . If

$$(11.4.1) \quad \sum_{x \in X} f(x) < \infty,$$

then f is said to be *summable* on X . Of course, if f has finite support in X , then f is summable on X . Suppose now that f is summable on X , and let us check that f vanishes at infinity on X , as in Section 2.5.

11.4.1 Summability and vanishing at infinity

Let $\epsilon > 0$ be given, and put

$$(11.4.2) \quad E_\epsilon(f) = \{x \in X : f(x) \geq \epsilon\}.$$

If A is a nonempty finite subset of $E_\epsilon(f)$, then

$$(11.4.3) \quad \epsilon(\#A) \leq \sum_{x \in A} f(x) \leq \sum_{x \in X} f(x),$$

where $\#A$ denotes the number of elements of A . Thus

$$(11.4.4) \quad \#A \leq (1/\epsilon) \sum_{x \in X} f(x).$$

It follows that $E_\epsilon(f)$ has only finitely many elements, with

$$(11.4.5) \quad \#E_\epsilon(f) \leq (1/\epsilon) \sum_{x \in X} f(x).$$

11.4.2 Another property of summable functions

Let us continue to suppose that f is summable on X , and let $\epsilon > 0$ be given again. Observe that there is a nonempty finite subset $A(\epsilon)$ of X such that

$$(11.4.6) \quad \sum_{x \in X} f(x) - \epsilon < \sum_{x \in A(\epsilon)} f(x),$$

by the definition (11.2.2) of the sum over X . If A is a finite subset of X that contains $A(\epsilon)$, then

$$(11.4.7) \quad \sum_{x \in A(\epsilon)} f(x) \leq \sum_{x \in A} f(x) \leq \sum_{x \in X} f(x) < \sum_{x \in A(\epsilon)} f(x) + \epsilon.$$

Let B be a nonempty finite subset of X that is disjoint from $A(\epsilon)$, so that

$$(11.4.8) \quad \sum_{x \in A(\epsilon)} f(x) + \sum_{x \in B} f(x) = \sum_{x \in A(\epsilon) \cup B} f(x) \leq \sum_{x \in X} f(x).$$

This implies that

$$(11.4.9) \quad \sum_{x \in B} f(x) \leq \sum_{x \in X} f(x) - \sum_{x \in A(\epsilon)} f(x).$$

If $A(\epsilon) \neq X$, then it follows that

$$(11.4.10) \quad \sum_{x \in X \setminus A(\epsilon)} f(x) \leq \sum_{x \in X} f(x) - \sum_{x \in A(\epsilon)} f(x) < \epsilon.$$

If $A(\epsilon) = X$, then the sum on the left side of (11.4.10) may be interpreted as being equal to 0.

11.4.3 More on nonnegative summable functions

If f is summable on X and E is a nonempty subset of X , then the restriction of f to E is summable on E . More precisely, the sum of $f(x)$ over $x \in E$ is less than or equal to the sum of $f(x)$ over $x \in X$, as in (11.3.4).

Similarly, let f and g be nonnegative real-valued functions on X such that $f(x) \leq g(x)$ for every $x \in X$. If g is summable on X , then f is summable on X too, by (11.3.11).

11.5 Some linearity properties

Let X be a nonempty set, and let f be a nonnegative extended real-valued function on X . Also let t be a positive real number, so that tf is a nonnegative extended real-valued function on X as well. If A is a nonempty finite subset of X , then

$$(11.5.1) \quad \sum_{x \in A} tf(x) = t \sum_{x \in A} f(x).$$

Using this, one can check that

$$(11.5.2) \quad \sum_{x \in X} tf(x) = t \sum_{x \in X} f(x).$$

More precisely, this can be obtained from (11.1.5) and the definition (11.2.2) of the sum over X .

11.5.1 An additivity property

Let g be another nonnegative extended real-valued function on X , so that $f+g$ is a nonnegative extended real-valued function on X too. If A is a nonempty

finite subset of X , then

$$(11.5.3) \quad \sum_{x \in A} (f(x) + g(x)) = \sum_{x \in A} f(x) + \sum_{x \in A} g(x).$$

One can use this to show that

$$(11.5.4) \quad \sum_{x \in X} (f(x) + g(x)) = \sum_{x \in X} f(x) + \sum_{x \in X} g(x).$$

More precisely, if A is a nonempty finite subset of X , then

$$(11.5.5) \quad \sum_{x \in A} (f(x) + g(x)) \leq \sum_{x \in X} f(x) + \sum_{x \in X} g(x),$$

by (11.5.3). This implies that

$$(11.5.6) \quad \sum_{x \in X} (f(x) + g(x)) \leq \sum_{x \in X} f(x) + \sum_{x \in X} g(x).$$

11.5.2 The other inequality

We would like to verify that

$$(11.5.7) \quad \sum_{x \in X} f(x) + \sum_{x \in X} g(x) \leq \sum_{x \in X} (f(x) + g(x)),$$

in order to get (11.5.4). This is trivial when the right side of (11.5.7) is $+\infty$, and so we may suppose that it is finite, so that $f + g$ is summable on X . It follows that f and g are summable on X , because $f, g \leq f + g$, as in the previous section. In particular, $f(x)$ and $g(x)$ are finite for every $x \in X$. Let A and B be nonempty finite subsets of X , and observe that

$$(11.5.8) \quad \begin{aligned} \sum_{x \in A} f(x) + \sum_{x \in B} g(x) &\leq \sum_{x \in A \cup B} f(x) + \sum_{x \in A \cup B} g(x) \\ &= \sum_{x \in A \cup B} (f(x) + g(x)) \leq \sum_{x \in X} (f(x) + g(x)). \end{aligned}$$

Thus

$$(11.5.9) \quad \sum_{x \in A} f(x) \leq \sum_{x \in X} (f(x) + g(x)) - \sum_{x \in B} g(x).$$

This implies that

$$(11.5.10) \quad \sum_{x \in X} f(x) \leq \sum_{x \in X} (f(x) + g(x)) - \sum_{x \in B} g(x)$$

for every nonempty finite set $B \subseteq X$. Equivalently,

$$(11.5.11) \quad \sum_{x \in B} g(x) \leq \sum_{x \in X} (f(x) + g(x)) - \sum_{x \in X} f(x)$$

for every nonempty finite set $B \subseteq X$, and hence

$$(11.5.12) \quad \sum_{x \in X} g(x) \leq \sum_{x \in X} (f(x) + g(x)) - \sum_{x \in X} f(x).$$

This shows that (11.5.7) holds under these conditions, as desired.

11.5.3 A related additivity property

Let f be a nonnegative extended real-valued function on X again, and let E_1, E_2 be disjoint nonempty subsets of X . It is easy to see that

$$(11.5.13) \quad \sum_{x \in E_1 \cup E_2} f(x) = \sum_{x \in E_1} f(x) + \sum_{x \in E_2} f(x),$$

using (11.5.4). More precisely, this can be obtained by expressing f on $E_1 \cup E_2$ as the sum of two functions, with supports contained in E_1 and E_2 .

11.6 ℓ^1 Spaces

A real or complex-valued function f on a nonempty set X is said to be *summable* on X if $|f|$ is summable as a nonnegative real-valued function on X . Let $\ell^1(X, \mathbf{R})$ and $\ell^1(X, \mathbf{C})$ be the spaces of real and complex-valued summable functions on X , respectively. Of course, if a real or complex-valued function f on X has finite support, then f is summable on X , so that

$$(11.6.1) \quad c_{00}(X, \mathbf{R}) \subseteq \ell^1(X, \mathbf{R}), \quad c_{00}(X, \mathbf{C}) \subseteq \ell^1(X, \mathbf{C}).$$

If f is any real or complex-valued summable function on X , then $|f|$ vanishes at infinity on X , as in Section 11.4. This means that f vanishes at infinity on X too, so that

$$(11.6.2) \quad \ell^1(X, \mathbf{R}) \subseteq c_0(X, \mathbf{R}), \quad \ell^1(X, \mathbf{C}) \subseteq c_0(X, \mathbf{C}).$$

11.6.1 The ℓ^1 norm

If f is a real or complex-valued summable function on X , then put

$$(11.6.3) \quad \|f\|_1 = \sum_{x \in X} |f(x)|.$$

This is a nonnegative real number, which is equal to 0 exactly when $f(x) = 0$ for every $x \in X$. If $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then $|tf| = |t||f|$ is summable on X , by (11.5.2), and we have that

$$(11.6.4) \quad \|tf\|_1 = \sum_{x \in X} |tf(x)| = |t| \sum_{x \in X} |f(x)| = |t| \|f\|_1.$$

Similarly, if g is another real or complex-valued summable function on X , as appropriate, then it is easy to see that $f + g$ is summable on X as well, with

$$(11.6.5) \quad \|f + g\|_1 = \sum_{x \in X} |f(x) + g(x)| \leq \sum_{x \in X} (|f(x)| + |g(x)|) = \|f\|_1 + \|g\|_1.$$

Thus $\ell^1(X, \mathbf{R})$ and $\ell^1(X, \mathbf{C})$ are linear subspaces of the real and complex vector spaces of all real and complex-valued functions on X , respectively, and (11.6.3) defines a norm on each of $\ell^1(X, \mathbf{R})$ and $\ell^1(X, \mathbf{C})$.

11.6.2 The ℓ^1 metric

One can verify that

$$(11.6.6) \quad d_1(f, g) = \|f - g\|_1$$

defines a metric on each of $\ell^1(X, \mathbf{R})$ and $\ell^1(X, \mathbf{C})$, using (11.6.4) and (11.6.5), as usual.

Remember that $\|f\|_\infty$ denotes the supremum norm of a bounded real or complex-valued f on X , as in (1.13.4). If f is a real or complex-valued summable function on X , then

$$(11.6.7) \quad |f(x)| \leq \|f\|_1$$

for every $x \in X$. This implies that f is bounded on X , with

$$(11.6.8) \quad \|f\|_\infty \leq \|f\|_1.$$

If f, g are real or complex-valued summable functions on X , then it follows that

$$(11.6.9) \quad d_\infty(f, g) \leq d_1(f, g),$$

where $d_\infty(f, g)$ is the supremum metric, as in (1.13.8).

11.6.3 Density of $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ in $\ell^1(X, \mathbf{R})$, $\ell^1(X, \mathbf{C})$

Let us check that $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ are dense in $\ell^1(X, \mathbf{R})$, $\ell^1(X, \mathbf{C})$, respectively, with respect to the ℓ^1 metric (11.6.6). Of course, this is trivial when X has only finitely many elements, and so we suppose that X is an infinite set.

Let f be a real or complex-valued summable function on X , and let $\epsilon > 0$ be given. Remember that there is a finite subset $A(\epsilon)$ of X such that

$$(11.6.10) \quad \sum_{x \in X \setminus A(\epsilon)} |f(x)| < \epsilon,$$

as in (11.4.10). Let f_ϵ be the real or complex-valued function, as appropriate, defined on X by

$$(11.6.11) \quad \begin{aligned} f_\epsilon(x) &= f(x) && \text{when } x \in A(\epsilon) \\ &= 0 && \text{when } x \in X \setminus A(\epsilon). \end{aligned}$$

Thus f_ϵ has finite support in X . Note that $f - f_\epsilon$ is equal to 0 on $A(\epsilon)$, and to f on $X \setminus A(\epsilon)$. This implies that

$$(11.6.12) \quad \|f - f_\epsilon\|_1 = \sum_{x \in X \setminus A(\epsilon)} |f(x)| < \epsilon,$$

as desired.

11.7 Real-valued summable functions

Let f be a real-valued function on a nonempty set X , and put

$$(11.7.1) \quad f_+(x) = \max(f(x), 0) \quad \text{and} \quad f_-(x) = \max(-f(x), 0)$$

for each $x \in X$. These are nonnegative real-valued functions on X that satisfy

$$(11.7.2) \quad f_+(x) + f_-(x) = |f(x)|$$

and

$$(11.7.3) \quad f_+(x) - f_-(x) = f(x)$$

for every $x \in X$.

11.7.1 Summing real-valued summable functions

If f is summable on X , so that $|f|$ is summable on X , then f_+ and f_- are summable as nonnegative real-valued functions on X . In this case, we put

$$(11.7.4) \quad \sum_{x \in X} f(x) = \sum_{x \in X} f_+(x) - \sum_{x \in X} f_-(x).$$

Note that

$$(11.7.5) \quad \left| \sum_{x \in X} f(x) \right| \leq \sum_{x \in X} f_+(x) + \sum_{x \in X} f_-(x) = \sum_{x \in X} |f(x)|.$$

11.7.2 A property of the sum

Let f_1, f_2 be nonnegative real-valued summable functions on X such that

$$(11.7.6) \quad f(x) = f_1(x) - f_2(x)$$

for every $x \in X$. This implies that

$$(11.7.7) \quad f_1(x) + f_-(x) = f_+(x) + f_2(x)$$

for every $x \in X$, because of (11.7.3). Hence

$$(11.7.8) \quad \sum_{x \in X} f_1(x) + \sum_{x \in X} f_-(x) = \sum_{x \in X} f_+(x) + \sum_{x \in X} f_2(x),$$

as in (11.5.7).

It follows that

$$(11.7.9) \quad \sum_{x \in X} f_1(x) - \sum_{x \in X} f_2(x) = \sum_{x \in X} f_+(x) - \sum_{x \in X} f_-(x),$$

because the individual sums are all finite. This means that

$$(11.7.10) \quad \sum_{x \in X} f(x) = \sum_{x \in X} f_1(x) - \sum_{x \in X} f_2(x),$$

by (11.7.4).

11.7.3 Homogeneity of the sum

If f is a real-valued summable function on X and $t \in \mathbf{R}$, then tf is summable on X too, as in the previous section. One can check that

$$(11.7.11) \quad \sum_{x \in X} tf(x) = t \sum_{x \in X} f(x).$$

Of course, this is trivial when $t = 0$. If $t > 0$, then (11.7.11) can be derived from the analogous fact (11.5.2) for nonnegative real-valued functions on X , and the definition (11.7.4) of the sum for real-valued summable functions on X . If $t = -1$, then (11.7.11) can be obtained directly from (11.7.4).

11.7.4 Additivity of the sum

If f, g are real-valued summable functions on X , then $f + g$ is summable on X as well, as in the previous section again. One can verify that

$$(11.7.12) \quad \sum_{x \in X} (f(x) + g(x)) = \sum_{x \in X} f(x) + \sum_{x \in X} g(x).$$

We have already seen in Subsection 11.5.1 that this holds when f and g are nonnegative.

Otherwise, f and g can be expressed as differences of nonnegative real-valued summable functions on X , as in (11.7.3). This leads to an expression of $f + g$ as a difference of nonnegative real-valued summable functions on X . Each of the sums of f , g , and $f + g$ over X can be given as the difference of the corresponding sums of nonnegative summable functions, as in (11.7.4) and (11.7.10). Using this, one can reduce (11.7.12) to the analogous statement for nonnegative functions.

11.8 Complex-valued summable functions

Let f be a complex-valued summable function on a nonempty set X , and let $\operatorname{Re} f(x)$, $\operatorname{Im} f(x)$ be the real and imaginary parts of $f(x)$ for each $x \in X$, as

usual. Note that

$$(11.8.1) \quad |\operatorname{Re} f(x)|, |\operatorname{Im} f(x)| \leq |f(x)| \leq |\operatorname{Re} f(x)| + |\operatorname{Im} f(x)|$$

for every $x \in X$.

11.8.1 Summing complex-valued summable functions

This implies that f is summable as a complex-valued function on X if and only if its real and imaginary parts are summable as real-valued functions on X . In this case, we put

$$(11.8.2) \quad \sum_{x \in X} f(x) = \sum_{x \in X} \operatorname{Re} f(x) + i \sum_{x \in X} \operatorname{Im} f(x),$$

where the sums on the right side are defined as in the preceding section.

11.8.2 Additivity for complex-valued functions

If f, g are complex-valued summable functions on X , then $f + g$ is summable on X too, as in Subsection 11.6.1. It is easy to see that

$$(11.8.3) \quad \sum_{x \in X} (f(x) + g(x)) = \sum_{x \in X} f(x) + \sum_{x \in X} g(x).$$

Of course, the real and imaginary parts of $f + g$ are the same as the sums of the real and imaginary parts of f and g , respectively. Thus (11.8.3) follows from the definition (11.8.2) of the sum in the complex case and the analogous statement for the sum in the real case.

11.8.3 Homogeneity using complex numbers

If f is a complex-valued summable function on X and $t \in \mathbf{C}$, then tf is summable on X as well, as in Subsection 11.6.1 again. Let us check that

$$(11.8.4) \quad \sum_{x \in X} t f(x) = t \sum_{x \in X} f(x).$$

If $t \in \mathbf{R}$, then this can be obtained from the definition (11.8.2) of the sum in the complex case and the analogous property of the sum in the real case.

Similarly, if t is imaginary, then (11.8.4) can be reduced to the analogous statement in the real case. If $t \in \mathbf{C}$, then (11.8.4) can be derived from the previous two cases, using (11.8.3).

11.8.4 The absolute value of the sum in the complex case

Let f be a complex-valued summable function on X again, and observe that

$$(11.8.5) \quad \left| \sum_{x \in X} f(x) \right| \leq \left| \sum_{x \in X} \operatorname{Re} f(x) \right| + \left| \sum_{x \in X} \operatorname{Im} f(x) \right|,$$

by the definition (11.8.2) of the sum. This implies that

$$(11.8.6) \quad \left| \sum_{x \in X} f(x) \right| \leq \sum_{x \in X} |\operatorname{Re} f(x)| + \sum_{x \in X} |\operatorname{Im} f(x)|,$$

because of (11.7.5). It follows that

$$(11.8.7) \quad \left| \sum_{x \in X} f(x) \right| \leq 2 \sum_{x \in X} |f(x)|,$$

using the first step in (11.8.1).

Of course, we would rather have that

$$(11.8.8) \quad \left| \sum_{x \in X} f(x) \right| \leq \sum_{x \in X} |f(x)|.$$

This corresponds to the ordinary triangle inequality for the standard absolute value function on \mathbf{C} in the case of finite sums. In this situation, one can get (11.8.8) by approximating by finite sums. Some more details about this type of approximation will be given in the next section.

11.9 Some properties of the sum

Let f be a real or complex-valued function on a nonempty set X , and suppose for the moment that f has finite support in X . In this case, the sum $\sum_{x \in X} f(x)$ can be defined as a real or complex number by reducing to a finite sum. If f is a nonnegative real-valued function on X , then we have seen that the definition of the sum in Subsection 11.2.1 reduces to the same finite sum. If f is a real or complex-valued function on X , then we can apply this to $|f(x)|$, to get that f is summable on X . If f is a real-valued function on X , then the functions defined on X in (11.7.1) have finite support. This implies that the definition of the sum in Subsection 11.7.1 is the same as the finite sum, because of the analogous statement for nonnegative real-valued functions on X with finite support. Similarly, if f is a complex-valued function on X , then the real and imaginary parts of f have finite support in X . It follows that the definition of the sum in the preceding section is the same as the finite sum, because of the analogous statement for real-valued functions on X with finite support. In particular, (11.8.8) holds, because of the ordinary triangle inequality for finite sums.

11.9.1 Continuity of the sum on $\ell^1(X, \mathbf{R})$, $\ell^1(X, \mathbf{C})$

Now let f, g be real or complex-valued summable functions on X . Observe that

$$(11.9.1) \quad \left| \sum_{x \in X} f(x) - \sum_{x \in X} g(x) \right| = \left| \sum_{x \in X} (f(x) - g(x)) \right| \leq 2 \sum_{x \in X} |f(x) - g(x)|.$$

This uses the linearity of the sum in the first step, and (11.8.7) in the second step. It follows that the mapping from a real or complex-valued summable function f on X to its sum $\sum_{x \in X} f(x)$ is uniformly continuous as a mapping from $\ell^1(X, \mathbf{R})$, $\ell^1(X, \mathbf{C})$ into \mathbf{R} , \mathbf{C} , respectively. Here we use the corresponding ℓ^1 metric on the domain, as in (11.6.6), and the standard Euclidean metric on the range.

One can use these properties of the sum to check that (11.8.8) holds for all complex-valued summable functions f on X . More precisely, if f has finite support in X , then we have already seen that (11.8.8) holds. Otherwise, if f is any complex-valued summable function on X , then f can be approximated by complex-valued functions on X with finite support with respect to the ℓ^1 metric, as in Subsection 11.6.3. To get that f satisfies (11.8.8) as well, one can use the continuity condition (11.9.1).

Once we have that (11.8.8) holds for every complex-valued summable function on X , we get that

$$(11.9.2) \quad \left| \sum_{x \in X} f(x) - \sum_{x \in X} g(x) \right| \leq \sum_{x \in X} |f(x) - g(x)|$$

for all complex-valued summable functions f, g on X . This is basically the same as (11.9.1), except that we use (11.8.8) in the second step. Of course, if f and g are real-valued summable functions on X , then we could already get this using (11.7.5).

11.9.2 More on sums over subsets

Let f be a real or complex-valued summable function on X again. If E is a nonempty subset of X , then the restriction of f to E is summable as a real or complex-valued function on E , as appropriate. This follows from the analogous statement for nonnegative real-valued summable functions, which was mentioned in Subsection 11.4.3. In particular, this means that $\sum_{x \in E} f(x)$ can be defined as a real or complex number, as appropriate, as in the previous two sections.

Note that

$$(11.9.3) \quad \sum_{x \in E} f(x) = \sum_{x \in X} f(x)$$

when $f(x) = 0$ for every $x \in X \setminus E$. This was mentioned in Subsection 11.3.1 when f is a nonnegative real-valued function on X . If f is a real-valued summable function on X with support contained in E , then the functions defined in (11.7.1) have support contained in E too. In this case, (11.9.3) can be reduced to the corresponding statement for nonnegative real-valued functions, because of the way that the sum was defined in Subsection 11.7.1. Similarly, if f is a complex-valued summable function on X with support contained in E , then (11.9.3) can be obtained from the corresponding statement for real-valued functions, applied to the real and imaginary parts of f .

11.9.3 Another additivity property

If E_1 and E_2 are nonempty disjoint subsets of X , then

$$(11.9.4) \quad \sum_{x \in E_1 \cup E_2} f(x) = \sum_{x \in E_1} f(x) + \sum_{x \in E_2} f(x).$$

This was mentioned in Subsection 11.5.3 when f is nonnegative. If f is any real or complex-valued summable function on X , then (11.9.4) can be obtained from the linearity of the sum, by expressing f on $E_1 \cup E_2$ as a sum of functions supported in E_1 and E_2 , as before.

Alternatively, if f is a real-valued summable function on X , then one can reduce to the case of nonnegative real-valued summable functions on X , because of the way that the sum is defined in Subsection 11.7.1. Similarly, if f is a complex-valued summable function on X , then one can apply the previous statement to the real and imaginary parts of f .

11.10 Generalized convergence

Let f be a real or complex-valued function on a nonempty set X . The sum

$$(11.10.1) \quad \sum_{x \in X} f(x)$$

is said to *converge in the generalized sense* if there is a real or complex number λ , as appropriate, such that for every $\epsilon > 0$ there is a nonempty finite subset A_ϵ of X with the property that

$$(11.10.2) \quad \left| \sum_{x \in A} f(x) - \lambda \right| < \epsilon$$

for every nonempty finite subset A of X with $A_\epsilon \subseteq A$. In this case, the value of the sum (11.10.1) is defined to be λ .

More precisely, one can check that λ is unique when it exists. Of course, if X has only finitely many elements, then one can take $A_\epsilon = X$ for every $\epsilon > 0$, to get that this condition holds trivially with λ equal to the finite sum (11.10.1).

11.10.1 Generalized convergence and summable functions

Suppose that f is summable on X , and let us check that the sum (11.10.1) converges in the generalized sense, with the same value of the sum as defined in Subsections 11.2.1 and 11.7.1, and 11.8.1. This is trivial when X has only finitely many elements, and so we suppose now that X is an infinite set.

If A is a nonempty finite subset of X , then

$$(11.10.3) \quad \sum_{x \in X} f(x) = \sum_{x \in A} f(x) + \sum_{x \in X \setminus A} f(x),$$

as in (11.9.4). Note that $X \setminus A \neq \emptyset$ in this situation, and that the restriction of f to $X \setminus A$ is summable, as in Subsection 11.9.2. It follows that

$$(11.10.4) \quad \left| \sum_{x \in A} f(x) - \sum_{x \in X} f(x) \right| = \left| \sum_{x \in X \setminus A} f(x) \right| \leq \sum_{x \in X \setminus A} |f(x)|,$$

using (11.7.5) or (11.8.8) in the second step, as appropriate.

Remember that for each $\epsilon > 0$ there is a nonempty finite subset $A(\epsilon)$ of X such that

$$(11.10.5) \quad \sum_{x \in X \setminus A(\epsilon)} |f(x)| < \epsilon,$$

as in (11.4.10). If $A(\epsilon) \subseteq A$, then $X \setminus A \subseteq X \setminus A(\epsilon)$, and hence

$$(11.10.6) \quad \sum_{x \in X \setminus A} |f(x)| \leq \sum_{x \in X \setminus A(\epsilon)} |f(x)|,$$

as in (11.3.4). Thus

$$(11.10.7) \quad \left| \sum_{x \in A} f(x) - \sum_{x \in X} f(x) \right| \leq \sum_{x \in X \setminus A} |f(x)| \leq \sum_{x \in X \setminus A(\epsilon)} |f(x)| < \epsilon$$

when A is a nonempty finite subset of X with $A(\epsilon) \subseteq A$, as desired.

Alternatively, if f is a nonnegative real-valued summable function on X , then the convergence of the sum (11.10.1) in the generalized sense can be obtained from (11.4.7). If f is a real-valued summable function on X , then f can be expressed as a difference of nonnegative real-valued summable functions on X , as in (11.7.3). In this case, the convergence of the sum (11.10.1) in the generalized sense can be reduced to the previous statement. Similarly, if f is a complex-valued summable function on X , then the real and imaginary parts of f are summable on X too. This permits one to reduce the convergence of (11.10.1) in the generalized sense to the analogous statements for the real and imaginary parts of f .

11.10.2 Getting summability

Let f be a real or complex-valued function on X again, and suppose that the sum (11.10.1) converges in the generalized sense. Applying the earlier definition with $\epsilon = 1$, we get that there is a real or complex number λ , as appropriate, and a nonempty finite subset A_1 of X such that

$$(11.10.8) \quad \left| \sum_{x \in A_1} f(x) - \lambda \right| < 1$$

for every nonempty finite subset A of X with $A_1 \subseteq A$.

In particular,

$$(11.10.9) \quad \left| \sum_{x \in A} f(x) \right| < |\lambda| + 1$$

for every nonempty finite subset A of X with $A_1 \subseteq A$. Let us check that

$$(11.10.10) \quad \left| \sum_{x \in B} f(x) \right| \leq |\lambda| + 1 + \sum_{x \in A_1} |f(x)|$$

for every nonempty finite subset B of X .

Of course, (11.10.10) follows from (11.10.9) when $A_1 \subseteq B$. Otherwise, we can apply (11.10.9) to $A = B \cup A_1$, to get that

$$(11.10.11) \quad \left| \sum_{x \in B \cup A_1} f(x) \right| < |\lambda| + 1.$$

This implies that

$$(11.10.12) \quad \begin{aligned} \left| \sum_{x \in B} f(x) \right| &= \left| \sum_{x \in B \cup A_1} f(x) - \sum_{x \in A_1 \setminus B} f(x) \right| \\ &\leq \left| \sum_{x \in B \cup A_1} f(x) \right| + \left| \sum_{x \in A_1 \setminus B} f(x) \right| \\ &\leq |\lambda| + 1 + \sum_{x \in A_1 \setminus B} |f(x)|. \end{aligned}$$

This shows that (11.10.10) also holds when $A_1 \setminus B \neq \emptyset$, as desired.

If f is a nonnegative real-valued function on X , then (11.10.10) implies that f is summable on X . If f is a real-valued function on X , then (11.10.10) implies that the positive and negative parts of f are summable on X . More precisely, one can get this by applying (11.10.10) to nonempty finite subsets B of X on which f has constant sign. It follows that f is summable on X too in this case. If f is a complex-valued function on X , then one can use (11.10.10) to get that the real and imaginary parts of f are summable on X , and hence that f is summable on X .

11.11 Compositions and sums

Let X and Y be nonempty sets, and let ϕ be a one-to-one mapping from X onto Y . Also let f be a real or complex-valued function on Y , so that $f(\phi(x))$ defines a real or complex-valued function on X , as appropriate. Note that

$$(11.11.1) \quad \sum_{x \in X} |f(\phi(x))| = \sum_{y \in Y} |f(y)|,$$

as in (11.3.2). Thus f is summable on Y if and only if $f(\phi(x))$ is summable on X .

11.11.1 Sums over X and Y

Let us check that

$$(11.11.2) \quad \sum_{x \in X} f(\phi(x)) = \sum_{y \in Y} f(y)$$

in this case. If f is a nonnegative real-valued function on Y , then (11.11.2) is the same as (11.11.1). If f is a real-valued function on Y , then one can get (11.11.2) by expressing f as a difference of nonnegative real-valued summable functions on Y , as in Section 11.7. If f is a complex-valued summable function on Y , then (11.11.2) can be obtained by applying the previous statement to the real and imaginary parts of f .

Similarly, if f is a real or complex-valued function on Y , then one can check directly that $\sum_{x \in X} f(\phi(x))$ converges in the generalized sense if and only if $\sum_{y \in Y} f(y)$ converges in the generalized sense, with the same value of the sums. This uses the fact that $A \mapsto \phi(A)$ defines a one-to-one correspondence between nonempty finite subsets of X and nonempty finite subsets of Y .

11.11.2 Countably-infinite sets X

Now let X be a countably infinite set, and let $\{x_j\}_{j=1}^{\infty}$ be a sequence of elements of X in which every element of X occurs exactly once. Thus $j \mapsto x_j$ is a one-to-one mapping from the set \mathbf{Z}_+ of positive integers onto X . Let f be a real or complex-valued function on X , so that $f(x_j)$ may be considered as a real or complex-valued function of $j \in \mathbf{Z}_+$, as appropriate. As before, f is summable on X if and only if $f(x_j)$ is summable on \mathbf{Z}_+ , in which case

$$(11.11.3) \quad \sum_{j \in \mathbf{Z}_+} f(x_j) = \sum_{x \in X} f(x).$$

Alternatively, the sum on the left converges in the generalized sense if and only if the sum on the right converges in the generalized sense, with the same value of the sum, as in the previous subsection.

11.11.3 Relation with infinite series

Remember that

$$(11.11.4) \quad \sum_{j=1}^{\infty} |f(x_j)| = \sum_{j \in \mathbf{Z}_+} |f(x_j)|,$$

as in (11.2.4). Thus $f(x_j)$ is summable on \mathbf{Z}_+ if and only if the infinite series on the left side of (11.11.4) converges in the usual sense, which means that $\sum_{j=1}^{\infty} f(x_j)$ converges absolutely.

Under these conditions,

$$(11.11.5) \quad \sum_{j=1}^{\infty} f(x_j) = \sum_{j \in \mathbf{Z}_+} f(x_j).$$

More precisely, (11.11.5) is the same as (11.11.4) when f is real-valued and nonnegative. If f is real-valued, then (11.11.5) can be obtained from the previous statement by expressing f as a difference of nonnegative real-valued summable functions. If f is complex-valued, then one can get (11.11.5) by considering the real and imaginary parts separately.

Alternatively, if the sum on the right side of (11.11.5) converges in the generalized sense, then it is easy to see that the sum on the left side of (11.11.5) converges, and with the same value of the sum.

11.12 Completeness of ℓ^1 spaces

Let X be a nonempty set, and let $\{f_j\}_{j=1}^\infty$ be a sequence of real or complex-valued functions on X that converge pointwise to a real or complex-valued function f on X , as appropriate. Suppose also that the f_j 's are summable on X , with bounded ℓ^1 norms, so that there is a nonnegative real number C such that

$$(11.12.1) \quad \sum_{x \in X} |f_j(x)| \leq C$$

for every $j \geq 1$. We would like to show that f is summable on X too under these conditions, with

$$(11.12.2) \quad \sum_{x \in X} |f(x)| \leq C.$$

If A is any nonempty finite subset of X , then

$$(11.12.3) \quad \sum_{x \in A} |f(x)| = \lim_{j \rightarrow \infty} \sum_{x \in A} |f_j(x)| \leq C,$$

using basic properties of limits in the first step, and (11.12.1) in the second step. This implies (11.12.2), as desired.

11.12.1 Cauchy sequences in $\ell^1(X, \mathbf{R})$, $\ell^1(X, \mathbf{C})$

Let us now check that $\ell^1(X, \mathbf{R})$ and $\ell^1(X, \mathbf{C})$ are complete as metric spaces, with respect to the ℓ^1 metric (11.6.6). Let $\{f_j\}_{j=1}^\infty$ be a sequence of real or complex-valued summable functions on X that is a Cauchy sequence with respect to the ℓ^1 metric. This means that for each $\epsilon > 0$ there is an $L(\epsilon) \in \mathbf{Z}_+$ such that

$$(11.12.4) \quad \sum_{x \in X} |f_j(x) - f_l(x)| = \|f_j - f_l\|_1 < \epsilon$$

for every $j, l \geq L(\epsilon)$.

In particular, if $x \in X$, then it follows that

$$(11.12.5) \quad |f_j(x) - f_l(x)| < \epsilon$$

for every $j, l \geq L(\epsilon)$. Thus $\{f_j(x)\}_{j=1}^\infty$ is a Cauchy sequence in \mathbf{R} or \mathbf{C} , as appropriate, with respect to the standard Euclidean metric.

11.12.2 Pointwise convergence

It is well known that \mathbf{R} and \mathbf{C} are complete as metric spaces with respect to the corresponding Euclidean metric. This means that $\{f_j(x)\}_{j=1}^\infty$ converges in \mathbf{R} or \mathbf{C} , as appropriate, for each $x \in X$. Put

$$(11.12.6) \quad f(x) = \lim_{j \rightarrow \infty} f_j(x)$$

for every $x \in X$, which defines f as a real or complex-valued function on X , as appropriate.

We would like to show that f is summable on X , and that $\{f_j\}_{j=1}^\infty$ converges to f with respect to the ℓ^1 metric. Of course, if X has only finitely many elements, then this follows easily from pointwise convergence.

11.12.3 Convergence in $\ell^1(X, \mathbf{R})$, $\ell^1(X, \mathbf{C})$

Let $\epsilon > 0$ and $l \geq L(\epsilon)$ be given, and consider $\{f_j - f_l\}_{j=l}^\infty$ as a sequence of summable functions on X that converges pointwise to $f - f_l$. Using (11.12.4) and the remarks at the beginning of the section, we get that $f - f_l$ is summable on X , with

$$(11.12.7) \quad \sum_{x \in X} |f(x) - f_l(x)| \leq \epsilon.$$

In particular, f is summable on X , because f_l is summable on X . More precisely, we can simply take $\epsilon = 1$ and $l = L(1)$ in this step.

It follows from (11.12.7) that $\{f_j\}_{j=1}^\infty$ converges to f with respect to the ℓ^1 metric, as desired.

11.13 Monotone convergence

Let X be a nonempty set, and let $\{f_j\}_{j=1}^\infty$ be a sequence of nonnegative real-valued functions on X . Suppose that the f_j 's are monotonically increasing in j , so that

$$(11.13.1) \quad f_j(x) \leq f_{j+1}(x)$$

for every $x \in X$ and $j \geq 1$. Put

$$(11.13.2) \quad f(x) = \sup_{j \geq 1} f_j(x)$$

for each $x \in X$, which defines f as a nonnegative extended real-valued function on X .

Equivalently, this means that

$$(11.13.3) \quad f_j(x) \rightarrow f(x) \quad \text{as } j \rightarrow \infty$$

for every $x \in X$, as in Subsection 11.1.4. Note that

$$(11.13.4) \quad \sum_{x \in X} f_j(x) \leq \sum_{x \in X} f_{j+1}(x)$$

for every $j \geq 1$, because of (11.13.1), as in (11.3.11).

Similarly,

$$(11.13.5) \quad \sum_{x \in X} f_j(x) \leq \sum_{x \in X} f(x)$$

for every $j \geq 1$, because $f_j(x) \leq f(x)$ for every $x \in X$ and $j \geq 1$, by construction. Thus

$$(11.13.6) \quad \sup_{j \geq 1} \left(\sum_{x \in X} f_j(x) \right) \leq \sum_{x \in X} f(x).$$

11.13.1 Convergence of the sums

We would like to show that

$$(11.13.7) \quad \sup_{j \geq 1} \left(\sum_{x \in X} f_j(x) \right) = \sum_{x \in X} f(x)$$

under these conditions. This is basically the same as saying that

$$(11.13.8) \quad \sum_{x \in X} f_j(x) \rightarrow \sum_{x \in X} f(x) \quad \text{as } j \rightarrow \infty,$$

because of (11.13.4).

More precisely, if f_j is summable on X for each $j \geq 1$, then (11.13.7) is equivalent to (11.13.8), as in Subsection 11.1.4. Otherwise, if f_j is not summable on X for some $j \geq 1$, then f_j is not summable on X for all sufficiently large j , because of (11.13.1). In this case, it is easy to see that f is not summable on X , so that (11.13.7) holds, and that (11.13.8) holds in a suitable sense.

11.13.2 The proof of convergence

In order to show (11.13.7) or (11.13.8), let a be a real number such that

$$(11.13.9) \quad a < \sum_{x \in X} f(x).$$

By definition of the sum on the right, there is a nonempty finite subset A of X such that

$$(11.13.10) \quad a < \sum_{x \in A} f(x).$$

Observe that

$$(11.13.11) \quad \sum_{x \in A} f_j(x) \rightarrow \sum_{x \in A} f(x) \quad \text{as } j \rightarrow \infty,$$

because of (11.13.3).

It follows that

$$(11.13.12) \quad a < \sum_{x \in A} f_j(x)$$

for all but finitely many $j \geq 1$. This implies that

$$(11.13.13) \quad a < \sum_{x \in X} f_j(x)$$

for the same j 's, and hence all but finitely many $j \geq 1$.

Using this, one can check that (11.13.7) or (11.13.8) holds, as desired. This is the *monotone convergence theorem* for sums. Of course, if X has only finitely many elements, then this is more elementary, as in (11.13.11).

11.14 Dominated convergence

Let X be a nonempty set, and let $\{f_j\}_{j=1}^{\infty}$ be a sequence of real or complex-valued functions on X that converges pointwise to a real or complex-valued function f on X , as appropriate. Suppose that there is a nonnegative real-valued summable function g on X such that

$$(11.14.1) \quad |f_j(x)| \leq g(x)$$

for every $x \in X$ and $j \geq 1$. This implies that

$$(11.14.2) \quad |f(x)| \leq g(x)$$

for every $x \in X$, because $\{f_j(x)\}_{j=1}^{\infty}$ converges to $f(x)$, by hypothesis.

In particular, f_j is summable on X for each $j \geq 1$, because of (11.14.1), and f is summable on X too, because of (11.14.2). We would like to show that

$$(11.14.3) \quad \lim_{j \rightarrow \infty} \sum_{x \in X} |f_j(x) - f(x)| = 0$$

under these conditions.

Of course, if X has only finitely many elements, then this follows from standard results about sums of convergent sequences. Thus we may suppose that X has infinitely many elements.

11.14.1 Getting some sums to be uniformly small

Let $\epsilon > 0$ be given. Because g is summable on X , there is a nonempty finite subset A_0 of X such that

$$(11.14.4) \quad \sum_{x \in X \setminus A_0} g(x) < \epsilon/3,$$

as in (11.4.10). Note that

$$(11.14.5) \quad |f_j(x) - f(x)| \leq |f_j(x)| + |f(x)| \leq 2g(x)$$

for every $x \in X$ and $j \geq 1$, by (11.14.1) and (11.14.2). It follows that

$$(11.14.6) \quad \sum_{x \in X \setminus A_0} |f_j(x) - f(x)| \leq \sum_{x \in X \setminus A_0} 2g(x) < 2\epsilon/3$$

for every $j \geq 1$.

11.14.2 Getting some sums to be small when j is large

As before,

$$(11.14.7) \quad \lim_{j \rightarrow \infty} \sum_{x \in A_0} |f_j(x) - f(x)| = 0,$$

because A_0 has only finitely many elements. Hence there is an $L \in \mathbf{Z}_+$ such that

$$(11.14.8) \quad \sum_{x \in A_0} |f_j(x) - f(x)| < \epsilon/3$$

for every $j \geq L$. Combining this with (11.14.6), we get that

$$(11.14.9) \quad \begin{aligned} \sum_{x \in X} |f_j(x) - f(x)| &= \sum_{x \in A_0} |f_j(x) - f(x)| + \sum_{x \in X \setminus A_0} |f_j(x) - f(x)| \\ &< \epsilon/3 + 2\epsilon/3 = \epsilon \end{aligned}$$

for every $j \geq L$, as desired.

11.14.3 A simple corollary

Observe that

$$(11.14.10) \quad \begin{aligned} \left| \sum_{x \in X} f_j(x) - \sum_{x \in X} f(x) \right| &= \left| \sum_{x \in X} (f_j(x) - f(x)) \right| \\ &\leq \sum_{x \in X} |f_j(x) - f(x)| \end{aligned}$$

for every $j \geq 1$. This uses (11.7.5) or (11.8.8), as appropriate, in the second step. Thus (11.14.3) implies that

$$(11.14.11) \quad \lim_{j \rightarrow \infty} \sum_{x \in X} f_j(x) = \sum_{x \in X} f(x).$$

The fact that this holds under these conditions is the *dominated convergence theorem* for sums. If X has only finitely many elements, then this is a standard result about sums of convergent sequences, as before.

11.15 Nonnegative sums of sums

Let X be a nonempty set, and let f be a nonnegative real-valued function on X . Also let I be a nonempty set, and let E_j be a nonempty subset of X for each $j \in I$. Suppose that

$$(11.15.1) \quad E_j \cap E_l = \emptyset$$

for every $j, l \in I$ with $j \neq l$, and put

$$(11.15.2) \quad E = \bigcup_{j \in I} E_j.$$

11.15.1 Sums over the E_j 's

As usual,

$$(11.15.3) \quad \sum_{x \in E_j} f(x)$$

is defined as a nonnegative extended real number for each $j \in I$, so that (11.15.3) may be considered as a nonnegative extended real-valued function of j on I .

Thus

$$(11.15.4) \quad \sum_{j \in I} \left(\sum_{x \in E_j} f(x) \right)$$

can also be defined as a nonnegative extended real number, as in Subsection 11.2.4. Of course,

$$(11.15.5) \quad \sum_{x \in E} f(x)$$

can be defined as a nonnegative extended real number as well.

We would like to show that (11.15.4) is equal to (11.3.3). This is elementary when E is a finite set, which is the same as saying that I is a finite set, and that E_j is a finite set for each $j \in I$.

11.15.2 Estimating the sum over E

Let us first verify that (11.15.5) is less than or equal to (11.15.4). Let A be a nonempty finite subset of E . Put

$$(11.15.6) \quad I_A = \{j \in I : A \cap E_j \neq \emptyset\},$$

which is a nonempty finite subset of I , and observe that

$$(11.15.7) \quad A = \bigcup_{j \in I_A} A \cap E_j.$$

Thus

$$(11.15.8) \quad \sum_{x \in A} f(x) = \sum_{j \in I_A} \left(\sum_{x \in A \cap E_j} f(x) \right),$$

as mentioned earlier.

If $j \in I_A$, then

$$(11.15.9) \quad \sum_{x \in A \cap E_j} f(x) \leq \sum_{x \in E_j} f(x),$$

because $A \cap E_j \subseteq E_j$. Combining this with (11.15.8), we get that

$$(11.15.10) \quad \sum_{x \in A} f(x) \leq \sum_{j \in I_A} \left(\sum_{x \in E_j} f(x) \right).$$

It follows that

$$(11.15.11) \quad \sum_{x \in A} f(x) \leq \sum_{j \in I} \left(\sum_{x \in E_j} f(x) \right),$$

because $I_A \subseteq I$. This implies that

$$(11.15.12) \quad \sum_{x \in E} f(x) \leq \sum_{j \in I} \left(\sum_{x \in E_j} f(x) \right),$$

by the definition of the sum over E on the left.

11.15.3 Estimating the iterated sum

Now let us check that (11.15.4) is less than or equal to (11.15.5). Let I_0 be a nonempty finite subset of I , and put

$$(11.15.13) \quad E(I_0) = \bigcup_{j \in I_0} E_j.$$

Observe that

$$(11.15.14) \quad \sum_{j \in I_0} \left(\sum_{x \in E_j} f(x) \right) = \sum_{x \in E(I_0)} f(x).$$

This follows from the analogous statement (11.5.13) for the union of two disjoint nonempty subsets of X .

Hence

$$(11.15.15) \quad \sum_{j \in I_0} \left(\sum_{x \in E_j} f(x) \right) \leq \sum_{x \in E} f(x),$$

because $E(I_0) \subseteq E$. This implies that

$$(11.15.16) \quad \sum_{j \in I} \left(\sum_{x \in E_j} f(x) \right) \leq \sum_{x \in E} f(x),$$

by the definition of the sum over I on the left. Combining (11.15.12) and (11.15.16), we get that (11.15.4) is equal to (11.15.5), as desired.

11.16 Real and complex sums

Let X be a nonempty set again, and let f be a real or complex-valued summable function on X . As in the previous section, we let I be a nonempty set, and we let E_j be a nonempty subset of X for each $j \in I$. We suppose that the E_j 's are pairwise disjoint, as in (11.15.1), and we let E be their union, as in (11.15.2).

The sum (11.15.3) of $f(x)$ over $x \in E_j$ is now defined as a real or complex number for each $j \in I$, as appropriate. Similarly, the sum (11.15.5) of $f(x)$ over $x \in E$ is defined as a real or complex number, as appropriate.

11.16.1 Some summability conditions

If $j \in I$, then

$$(11.16.1) \quad \left| \sum_{x \in E_j} f(x) \right| \leq \sum_{x \in E_j} |f(x)|,$$

as in (11.7.5) or (11.8.8), as appropriate. Thus

$$(11.16.2) \quad \sum_{j \in I} \left| \sum_{x \in E_j} f(x) \right| \leq \sum_{j \in I} \left(\sum_{x \in E_j} |f(x)| \right),$$

where these sums over I are defined as nonnegative extended real numbers in the usual way.

The remarks in the previous section imply that

$$(11.16.3) \quad \sum_{j \in I} \left(\sum_{x \in E_j} |f(x)| \right) = \sum_{x \in E} |f(x)|.$$

We also have that

$$(11.16.4) \quad \sum_{x \in E} |f(x)| \leq \sum_{x \in X} |f(x)| < \infty,$$

because f is summable on X , by hypothesis. It follows that

$$(11.16.5) \quad \sum_{j \in I} \left| \sum_{x \in E_j} f(x) \right| < \infty.$$

11.16.2 Iterated sums

This means that the sum (11.15.3) of $f(x)$ over $x \in E_j$ is summable as a real or complex-valued function of $j \in I$, as appropriate. Hence the sum (11.15.4) of this sum over $j \in I$ can be defined as a real or complex number, as appropriate.

We would like to check that this sum is equal to the sum (11.15.5) of $f(x)$ over $x \in E$. If f is a nonnegative real-valued function on X , then this follows from the remarks in the previous section.

If f is a real-valued summable function on X , then f can be expressed as the difference of two nonnegative real-valued summable functions on X , as in (11.7.3). In this case, the equality of (11.15.4) and (11.15.5) follows from the analogous statement for nonnegative real-valued summable functions on X .

If f is a complex-valued summable function on X , then one can apply the previous statement to the real and imaginary parts of f .

11.16.3 Sums over Cartesian products

Let Y and Z be nonempty sets, and suppose that $X = Y \times Z$ is their Cartesian product. If $f(y, z)$ is a nonnegative real-valued function on $Y \times Z$, then

$$(11.16.6) \quad \sum_{z \in Z} f(y, z)$$

defines a nonnegative extended real-valued function of $y \in Y$, and

$$(11.16.7) \quad \sum_{y \in Y} f(y, z)$$

defines a nonnegative extended real-valued function of $z \in Z$. Thus

$$(11.16.8) \quad \sum_{y \in Y} \left(\sum_{z \in Z} f(y, z) \right)$$

and

$$(11.16.9) \quad \sum_{z \in Z} \left(\sum_{y \in Y} f(y, z) \right)$$

are defined as nonnegative extended real numbers, as is

$$(11.16.10) \quad \sum_{(y, z) \in Y \times Z} f(y, z).$$

The remarks in the previous section imply that (11.16.8) is equal to (11.16.10). This corresponds to expressing $Y \times Z$ as the union of the pairwise-disjoint nonempty subsets of the form $\{y\} \times Z$, with $y \in Y$.

Similarly, (11.16.9) is equal to (11.16.10), which corresponds to expressing $Y \times Z$ as the union of the pairwise-disjoint nonempty subsets of the form $Y \times \{z\}$, with $z \in Z$. In particular, it follows that (11.16.8) is equal to (11.16.9).

Suppose now that $f(y, z)$ is a real or complex-valued summable function on $Y \times Z$. In this case, (11.16.6) defines a real or complex-valued function of $y \in Y$, and (11.16.7) defines a real or complex-valued function of $z \in Z$, as appropriate. These functions are summable on Y and Z , respectively, as in (11.16.5). Hence (11.16.8) and (11.16.9) are defined as real or complex numbers, as appropriate. The sum (11.16.10) is also defined as a real or complex number, as appropriate, and it is equal to (11.16.8) and (11.16.9), as before.

11.17 Square-summable functions

Let X be a nonempty set, and let f be a real or complex-valued function on X . If $|f(x)|^2$ is summable on X , then we say that f is *square-summable* on X . In this case, we put

$$(11.17.1) \quad \|f\|_2 = \left(\sum_{x \in X} |f(x)|^2 \right)^{1/2},$$

using the nonnegative square root on the right side, as usual.

If $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then tf is square-summable on X too, because $|tf(x)|^2 = |t|^2 |f(x)|^2$ is summable on X . We also have that

$$(11.17.2) \quad \|tf\|_2 = \left(\sum_{x \in X} |t|^2 |f(x)|^2 \right)^{1/2} = |t| \|f\|_2.$$

11.17.1 The Cauchy–Schwarz inequality in this setting

Remember that

$$(11.17.3) \quad ab \leq \frac{1}{2}(a^2 + b^2)$$

for all nonnegative real numbers a, b , as in (2.3.6). Let g be another real or complex-valued square-summable function on X , as appropriate. Observe that

$$(11.17.4) \quad |f(x)| |g(x)| \leq \frac{1}{2} (|f(x)|^2 + |g(x)|^2)$$

for every $x \in X$, by (11.17.3).

This implies that

$$(11.17.5) \quad \begin{aligned} \sum_{x \in X} |f(x)| |g(x)| &\leq \sum_{x \in X} \frac{1}{2} (|f(x)|^2 + |g(x)|^2) \\ &= \frac{1}{2} \sum_{x \in X} |f(x)|^2 + \frac{1}{2} \sum_{x \in X} |g(x)|^2 = \frac{1}{2} \|f\|_2^2 + \frac{1}{2} \|g\|_2^2. \end{aligned}$$

In particular, $|f(x)| |g(x)|$ is summable on X when f and g are square-summable on X .

More precisely,

$$(11.17.6) \quad \sum_{x \in X} |f(x)| |g(x)| \leq \|f\|_2 \|g\|_2$$

when f and g are square-summable on X . This is another version of the Cauchy-Schwarz inequality. If $\|f\|_2 = \|g\|_2 = 1$, then (11.17.6) follows from (11.17.5). If $\|f\|_2, \|g\|_2 > 0$, then one can reduce to the previous case, using (11.17.2). Otherwise, if $f(x) = 0$ for every $x \in X$, or $g(x) = 0$ for every $x \in X$, then (11.17.6) is trivial.

11.17.2 The triangle inequality for $\|\cdot\|_2$

Note that

$$(11.17.7) \quad \begin{aligned} |f(x) + g(x)|^2 &\leq (|f(x)| + |g(x)|)^2 \\ &= |f(x)|^2 + 2|f(x)| |g(x)| + |g(x)|^2 \end{aligned}$$

for every $x \in X$. Hence

$$(11.17.8) \quad \begin{aligned} \sum_{x \in X} |f(x) + g(x)|^2 &\leq \sum_{x \in X} |f(x)|^2 + 2 \sum_{x \in X} |f(x)| |g(x)| + \sum_{x \in X} |g(x)|^2. \end{aligned}$$

If f and g are square-summable on X , then it follows that $f + g$ is square-summable on X , because $|f(x)| |g(x)|$ is summable on X , as before.

Using (11.17.6), we get that

$$(11.17.9) \quad \|f + g\|_2^2 \leq \|f\|_2^2 + 2\|f\|_2 \|g\|_2 + \|g\|_2^2 = (\|f\|_2 + \|g\|_2)^2.$$

Equivalently, this means that

$$(11.17.10) \quad \|f + g\|_2 \leq \|f\|_2 + \|g\|_2.$$

11.18 ℓ^2 Spaces

Let X be a nonempty set again, and let $\ell^2(X, \mathbf{R})$ and $\ell^2(X, \mathbf{C})$ be the spaces of real and complex-valued square-summable functions on X , respectively. These are linear subspaces of the real and complex vector spaces of all real and complex-valued functions on X , respectively, as in the previous section.

Note that (11.17.1) defines a norm on each of $\ell^2(X, \mathbf{R})$ and $\ell^2(X, \mathbf{C})$, because of (11.17.2), (11.17.10), and the fact that $\|f\|_2 = 0$ if and only if $f(x) = 0$ for every $x \in X$. This implies that

$$(11.18.1) \quad d_2(f, g) = \|f - g\|_2$$

defines a metric on each of $\ell^2(X, \mathbf{R})$ and $\ell^2(X, \mathbf{C})$, as usual.

11.18.1 Comparison with the supremum norm

If f is a real or complex-valued square-summable function on X , then

$$(11.18.2) \quad |f(x)| \leq \|f\|_2$$

for every $x \in X$, by the definition (11.17.1) of $\|f\|_2$. It follows that f is bounded on X , with

$$(11.18.3) \quad \|f\|_\infty \leq \|f\|_2.$$

Here $\|f\|_\infty$ denotes the supremum norm of f , as in (1.13.4). If g is another real or complex-valued square-summable function on X , as appropriate, then we get that

$$(11.18.4) \quad d_\infty(f, g) \leq d_2(f, g),$$

where $d_\infty(f, g)$ is the supremum metric, as in (1.13.8).

Let f be a real or complex-valued square-summable function on X again, so that $|f|^2$ is summable on X . This implies that $|f|^2$ vanishes at infinity on X , as in Section 11.4. Using this, it is easy to see that f vanishes at infinity on X as well. Thus

$$(11.18.5) \quad \ell^2(X, \mathbf{R}) \subseteq c_0(X, \mathbf{R}), \quad \ell^2(X, \mathbf{C}) \subseteq c_0(X, \mathbf{C}).$$

11.18.2 Comparison with the ℓ^1 norm

Let f be a real or complex-valued summable function on X , and remember that f is bounded on X , as in Subsection 11.6.2. Observe that

$$(11.18.6) \quad \sum_{x \in X} |f(x)|^2 \leq \|f\|_\infty \sum_{x \in X} |f(x)| = \|f\|_\infty \|f\|_1 \leq \|f\|_1^2,$$

where $\|f\|_1$ is the ℓ^1 norm of f , as in (11.6.3). This implies that f is square-summable on X , with

$$(11.18.7) \quad \|f\|_2 \leq \|f\|_1.$$

Hence

$$(11.18.8) \quad \ell^1(X, \mathbf{R}) \subseteq \ell^2(X, \mathbf{R}), \quad \ell^1(X, \mathbf{C}) \subseteq \ell^2(X, \mathbf{C}).$$

If g is another real or complex-valued summable function on X , as appropriate, then we have that

$$(11.18.9) \quad d_2(f, g) \leq d_1(f, g),$$

where $d_1(f, g)$ is the ℓ^1 metric, as in (11.6.6).

11.18.3 Density of $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ in $\ell^2(X, \mathbf{R})$, $\ell^2(X, \mathbf{C})$

In particular, real or complex-valued functions on X with finite support are square-summable. Let us verify that $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$ are dense in $\ell^2(X, \mathbf{R})$, $\ell^2(X, \mathbf{C})$, respectively, with respect to the ℓ^2 metric (11.18.1). This is very similar to the analogous argument for ℓ^1 spaces, in Subsection 11.6.3. As before, there is nothing to do when X has only finitely many elements, and so we suppose that X is an infinite set.

Let f be a real or complex-valued square-summable function on X , and let $\epsilon > 0$ be given. Because $|f|^2$ is summable on X , there is a finite subset $A(\epsilon)$ of X such that

$$(11.18.10) \quad \sum_{x \in X \setminus A(\epsilon)} |f(x)|^2 < \epsilon^2,$$

as in (11.4.10). Let f_ϵ be the real or complex-valued function on X , as appropriate, that is equal to f on $A(\epsilon)$ and to 0 on $X \setminus A(\epsilon)$. Thus f_ϵ has finite support in X , and $f - f_\epsilon$ is equal to 0 on $A(\epsilon)$, and to f on $X \setminus A(\epsilon)$. It follows that

$$(11.18.11) \quad \sum_{x \in X} |f(x) - f_\epsilon(x)|^2 = \sum_{x \in X \setminus A(\epsilon)} |f(x)|^2 < \epsilon^2,$$

so that

$$(11.18.12) \quad \|f - f_\epsilon\|_2 < \epsilon,$$

as desired.

11.19 Completeness of ℓ^2 spaces

Let X be a nonempty set, and let $\{f_j\}_{j=1}^\infty$ be a sequence of real or complex-valued functions on X that converges pointwise to a real or complex-valued function f on X , as appropriate. Suppose that the f_j 's are square-summable on X , with bounded ℓ^2 norms, so that

$$(11.19.1) \quad \|f_j\|_2 \leq C$$

for some nonnegative real number C and every $j \geq 1$. Equivalently, this means that

$$(11.19.2) \quad \sum_{x \in X} |f_j(x)|^2 \leq C^2$$

for every $j \geq 1$. Of course, $\{|f_j(x)|^2\}_{j=1}^\infty$ converges to $|f(x)|^2$ with respect to the standard Euclidean metric on \mathbf{R} for every $x \in X$, by well-known results about convergent sequences of real or complex numbers.

Under these conditions, f is a square-summable function on X , with

$$(11.19.3) \quad \sum_{x \in X} |f(x)|^2 \leq C^2.$$

This follows from the remarks at the beginning of Section 11.12, applied to $|f_j|^2$. Of course, (11.19.3) is the same as saying that

$$(11.19.4) \quad \|f\|_2 \leq C.$$

11.19.1 Cauchy sequences in $\ell^2(X, \mathbf{R})$, $\ell^2(X, \mathbf{C})$

We would like to show that $\ell^2(X, \mathbf{R})$, $\ell^2(X, \mathbf{C})$ are complete with respect to the ℓ^2 metric (11.18.1). Let $\{f_j\}_{j=1}^\infty$ be a sequence of real or complex-valued square-summable functions on X that is a Cauchy sequence with respect to the ℓ^2 metric. Thus for each $\epsilon > 0$ there is an $L(\epsilon) \in \mathbf{Z}_+$ such that

$$(11.19.5) \quad \|f_j - f_l\|_2 < \epsilon$$

for every $j, l \geq L(\epsilon)$. If $x \in X$, then it follows that

$$(11.19.6) \quad |f_j(x) - f_l(x)| < \epsilon$$

for every $j, l \geq L(\epsilon)$, by (11.18.2). This implies that $\{f_j(x)\}_{j=1}^\infty$ is a Cauchy sequence in \mathbf{R} or \mathbf{C} , as appropriate, with respect to the standard Euclidean metric.

Hence $\{f_j(x)\}_{j=1}^\infty$ converges in \mathbf{R} or \mathbf{C} , as appropriate, for every $x \in X$, because \mathbf{R} and \mathbf{C} are complete as metric spaces. Let $f(x)$ be the limit of this sequence for each $x \in X$, so that f defines a real or complex-valued function on X , as appropriate. We want to verify that f is square-summable on X , and that $\{f_j\}_{j=1}^\infty$ converges to f with respect to the ℓ^2 metric. If X has only finitely many elements, then f is automatically square-summable on X , and convergence with respect to the ℓ^2 metric can be obtained from pointwise convergence.

11.19.2 Convergence in $\ell^2(X, \mathbf{R})$, $\ell^2(X, \mathbf{C})$

Let $\epsilon > 0$ and $l \geq L(\epsilon)$ be given, and let us consider $\{f_j - f_l\}_{j=L(\epsilon)}^\infty$ as a sequence of square-summable functions on X that converges to $f - f_l$ pointwise on X . The remarks at the beginning of the section imply that $f - f_l$ is square-summable on X , with

$$(11.19.7) \quad \|f - f_l\|_2 \leq \epsilon,$$

because of (11.19.5). In particular, this holds with $\epsilon = 1$ and $l = L(1)$. This implies that f is square-summable on X , because $f_{L(1)}$ is square-summable on X . Using (11.19.7) again, we get that $\{f_j\}_{j=1}^\infty$ converges to f with respect to the ℓ^2 metric, as desired.

11.20 Inner products on ℓ^2 spaces

Let X be a nonempty set, and let f, g be square-summable real-valued functions on X . Under these conditions, $|f(x)| |g(x)|$ is summable as a nonnegative real-valued function on X , as in Subsection 11.17.1. This means that $f(x)g(x)$ is summable as a real-valued function on X , so that

$$(11.20.1) \quad \langle f, g \rangle = \langle f, g \rangle_{\ell^2(X, \mathbf{R})} = \sum_{x \in X} f(x) g(x)$$

can be defined as a real number, as in Subsection 11.7.1. This is the standard inner product on $\ell^2(X, \mathbf{R})$. In particular,

$$(11.20.2) \quad \langle f, f \rangle_{\ell^2(X, \mathbf{R})} = \sum_{x \in X} f(x)^2 = \|f\|_2^2,$$

where $\|f\|_2$ is the ℓ^2 norm of f , as in (11.17.1). As before,

$$(11.20.3) \quad |\langle f, g \rangle_{\ell^2(X, \mathbf{R})}| = \left| \sum_{x \in X} f(x) g(x) \right| \leq \sum_{x \in X} |f(x)| |g(x)| \leq \|f\|_2 \|g\|_2,$$

using the Cauchy-Schwarz inequality (11.17.6) in the third step. Of course, it is much easier to define this inner product on the space $c_{00}(X, \mathbf{R})$ of real-valued functions on X with finite support, by reducing the sum on the right side of (11.20.1) to a finite sum.

Note that

$$(11.20.4) \quad \langle f, g \rangle_{\ell^2(X, \mathbf{R})} = \sum_{x \in X} f(x) g(x) = \sum_{x \in X} g(x) f(x) = \langle g, f \rangle_{\ell^2(X, \mathbf{R})}.$$

If $t \in \mathbf{R}$, then tf is square-summable on X too, and

$$(11.20.5) \quad \begin{aligned} \langle tf, g \rangle_{\ell^2(X, \mathbf{R})} &= \sum_{x \in X} t f(x) g(x) \\ &= t \sum_{x \in X} f(x) g(x) = t \langle f, g \rangle_{\ell^2(X, \mathbf{R})}. \end{aligned}$$

If f_1 and f_2 are square-summable real-valued functions on X , then we have seen in Subsection 11.17.2 that $f_1 + f_2$ is square-summable on X as well. In this case, we have that

$$(11.20.6) \quad \begin{aligned} \langle f_1 + f_2, g \rangle_{\ell^2(X, \mathbf{R})} &= \sum_{x \in X} (f_1(x) + f_2(x)) g(x) \\ &= \sum_{x \in X} f_1(x) g(x) + \sum_{x \in X} f_2(x) g(x) \\ &= \langle f_1, g \rangle_{\ell^2(X, \mathbf{R})} + \langle f_2, g \rangle_{\ell^2(X, \mathbf{R})}, \end{aligned}$$

using the linearity of the sum in the second step. The analogous linearity properties of the inner product (11.20.1) in g can be obtained in the same way, or using (11.20.4).

11.20.1 An inner product on $\ell^2(X, \mathbf{C})$

Now let f, g be square-summable complex-valued functions on X . As before, $|f(x)||g(x)|$ is summable as a nonnegative real-valued function on X , so that $f(x)\overline{g(x)}$ is summable as a complex-valued function on X . Hence

$$(11.20.7) \quad \langle f, g \rangle = \langle f, g \rangle_{\ell^2(X, \mathbf{C})} = \sum_{x \in X} f(x) \overline{g(x)}$$

can be defined as a complex number, as in Subsection 11.8.1, which is the standard inner product on $\ell^2(X, \mathbf{C})$. As usual,

$$(11.20.8) \quad \langle f, f \rangle_{\ell^2(X, \mathbf{C})} = \sum_{x \in X} |f(x)|^2 = \|f\|_2^2,$$

where $\|f\|_2$ is the ℓ^2 norm of f . We also have that

$$(11.20.9) \quad |\langle f, g \rangle_{\ell^2(X, \mathbf{C})}| = \left| \sum_{x \in X} f(x) \overline{g(x)} \right| \leq \sum_{x \in X} |f(x)| |g(x)| \leq \|f\|_2 \|g\|_2,$$

using the Cauchy-Schwarz inequality (11.17.6) in the third step again.

In this situation, we have that

$$(11.20.10) \quad \overline{\langle f, g \rangle_{\ell^2(X, \mathbf{C})}} = \overline{\left(\sum_{x \in X} f(x) \overline{g(x)} \right)} = \sum_{x \in X} g(x) \overline{f(x)} = \langle g, f \rangle_{\ell^2(X, \mathbf{C})}.$$

If $t \in \mathbf{C}$, then tf is square-summable on X , and it is easy to see that

$$(11.20.11) \quad \langle tf, g \rangle_{\ell^2(X, \mathbf{C})} = t \langle f, g \rangle_{\ell^2(X, \mathbf{C})},$$

as before. Similarly, tg is square-summable on X , and

$$(11.20.12) \quad \langle f, tg \rangle_{\ell^2(X, \mathbf{C})} = \bar{t} \langle f, g \rangle_{\ell^2(X, \mathbf{C})},$$

by the same type of argument, or using (11.20.10). If f_1 and f_2 are square-summable complex-valued functions on X , then $f_1 + f_2$ are square-summable on X , and

$$(11.20.13) \quad \langle f_1 + f_2, g \rangle_{\ell^2(X, \mathbf{C})} = \langle f_1, g \rangle_{\ell^2(X, \mathbf{C})} + \langle f_2, g \rangle_{\ell^2(X, \mathbf{C})},$$

as in the real case. The analogous additivity property of the inner product (11.20.7) in g can be obtained in the same way, or using (11.20.10).

Chapter 12

Some additional topics

12.1 Lebesgue measure and integration

Let a, b be real numbers with $a < b$, and remember that

$$(12.1.1) \quad d_1(f, g) = \int_a^b |f(x) - g(x)| dx$$

defines a metric on the space $C([a, b], \mathbf{R})$ of continuous real-valued functions on $[a, b]$, as in Subsection 3.1.2. It is not difficult to see that this space is not complete with respect to (12.1.1), as in Section A.5. In order to get completeness, one can use the *Lebesgue integral*. Although we shall not discuss this in detail here, let us mention a few other topics related to Lebesgue's theory of measure and integration.

12.1.1 The bounded convergence theorem

Let $\{f_j\}_{j=1}^\infty$ be a sequence of continuous real-valued functions on $[a, b]$, and let f be another continuous real-valued function on $[a, b]$. If $\{f_j\}_{j=1}^\infty$ converges to f uniformly on $[a, b]$, then it is easy to see that $\{f_j\}_{j=1}^\infty$ converges to f with respect to (12.1.1). In particular, this implies that

$$(12.1.2) \quad \lim_{j \rightarrow \infty} \int_a^b f_j(x) dx = \int_a^b f(x) dx,$$

because

$$(12.1.3) \quad \left| \int_a^b f_j(x) dx - \int_a^b f(x) dx \right| \leq \int_a^b |f_j(x) - f(x)| dx$$

for every $j \geq 1$. However, one can give examples to show that this does not work if we only ask that $\{f_j\}_{j=1}^\infty$ converge to f pointwise on $[a, b]$, as in Section A.4.

Suppose now that $\{f_j\}_{j=1}^\infty$ is uniformly bounded on $[a, b]$, in the sense that there is a nonnegative real number A such that

$$(12.1.4) \quad |f_j(x)| \leq A$$

for every $x \in [a, b]$ and $j \geq 1$. If $\{f_j\}_{j=1}^\infty$ converges to f pointwise on $[a, b]$, then a classical theorem of Arzelà and Osgood implies that (12.1.2) holds, as in [54, 97, 165, 170, 243]. More precisely, this works when the f_j 's and f are Riemann integrable on $[a, b]$, instead of continuous. One can also get that $\{f_j\}_{j=1}^\infty$ converges to f with respect to (12.1.1), by considering $\{|f_j - f|\}_{j=1}^\infty$ in the previous statement. Questions like these can be treated more extensively using the Lebesgue integral.

12.1.2 Continuity almost everywhere

If f is a bounded real-valued function on $[a, b]$, then a famous theorem states that f is Riemann integrable on $[a, b]$ if and only if f is continuous at “almost every” point in $[a, b]$ with respect to Lebesgue measure. This means that the set of points in $[a, b]$ at which f is not continuous has Lebesgue measure equal to 0. This is discussed in many textbooks, as well as the articles [34, 198]. Some related results can be found in [164, 178], and also in [240], for Riemann–Stieltjes integrability.

12.1.3 Monotone functions

Now let f be a monotonically increasing real-valued function on $[a, b]$. It is well known that the one-sided limits of f exist at every point in (a, b) , as well as the appropriate one-sided limits at the endpoints a, b . This can be used to show that f is continuous at all but finitely or countably many points in $[a, b]$. A famous theorem states that f is also differentiable at almost every point in $[a, b]$, with respect to Lebesgue measure. Of course, the derivative is nonnegative when it exists, because of monotonicity.

One can show that

$$(12.1.5) \quad \int_a^b f'(x) dx \leq f(b) - f(a),$$

where the integral on the left is defined as a Lebesgue integral. If f has a jump discontinuity at any point in $[a, b]$, then the inequality in (12.1.5) is strict. There are also examples where f is both continuous and monotonically increasing on $[a, b]$, and the inequality in (12.1.5) is still strict.

12.1.4 Lipschitz functions

Suppose that f is a real-valued Lipschitz function on $[a, b]$, with respect to the standard Euclidean metric on \mathbf{R} and its restriction to $[a, b]$, but not necessarily monotonic. Another famous theorem states that f is differentiable almost

everywhere with respect to Lebesgue measure on $[a, b]$. It is easy to see that

$$(12.1.6) \quad |f'(x)| \leq \text{Lip}(f),$$

when the derivative exists. This permits the integral on the left side of (12.1.5) to be defined as a Lebesgue integral, and in fact one has

$$(12.1.7) \quad \int_a^b f'(x) dx = f(b) - f(a)$$

in this situation. If f is a real-valued Lipschitz function on \mathbf{R}^n for some $n \in \mathbf{Z}_+$, with respect to the standard Euclidean metrics on \mathbf{R}^n and \mathbf{R} , then another famous theorem states that f is differentiable almost everywhere with respect to n -dimensional Lebesgue measure on \mathbf{R}^n .

12.2 Banach spaces

We shall now suppose that the reader has some familiarity with abstract vector spaces and related notions. Let V be a vector space over the real or complex numbers. Thus V is a set, on which operations of addition and scalar multiplication over \mathbf{R} or \mathbf{C} have been defined. These operations should satisfy a number of standard conditions, such as associativity and commutativity of addition, and compatibility of scalar multiplication with addition on V and addition and multiplication on \mathbf{R} or \mathbf{C} , as appropriate. There should also be a distinguished additive element in V , denoted 0 .

12.2.1 Norms on vector spaces

As usual, a nonnegative real-valued function N on V is said to be a *norm* on V if it satisfies the following three conditions. First, $N(v) = 0$ if and only if $v = 0$. Second,

$$(12.2.1) \quad N(tv) = |t|N(v)$$

for every $v \in V$ and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate. Third,

$$(12.2.2) \quad N(v+w) \leq N(v) + N(w)$$

for every $v, w \in V$. In this case, it is easy to see that

$$(12.2.3) \quad d_N(v, w) = N(v - w)$$

defines a metric on V .

If V is also complete as a metric space with respect to (12.2.3), then V is said to be a *Banach space* with respect to N . Otherwise, one can pass to a suitable completion of V , but we shall get into that now.

12.2.2 Some remarks about convergent sequences

Suppose that $\{v_j\}_{j=1}^\infty, \{w_j\}_{j=1}^\infty$ are sequences of elements of V that converge to $v, w \in V$, respectively, with respect to (12.2.3). Under these conditions, one can check that $\{v_j + w_j\}_{j=1}^\infty$ converges to $v + w$ in V with respect to (12.2.3).

Similarly, suppose that $\{t_j\}_{j=1}^\infty$ is also a sequence of real or complex numbers that converges to a real or complex number t , with respect to the standard metric on \mathbf{R} or \mathbf{C} , as appropriate. In this case, one can verify that $\{t_j v_j\}_{j=1}^\infty$ converges to $t v$ in V with respect to (12.2.3). The proofs of these statements are analogous to those for the corresponding facts about sums and products of convergent sequences of real and complex numbers.

12.2.3 Linear subspaces

Let W be a *linear subspace* of V . This means that W is a subset of V that contains 0 and satisfies the following two properties. First, if $v, w \in W$, then

$$(12.2.4) \quad v + w \in W.$$

Second, if $v \in W$ and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then

$$(12.2.5) \quad t v \in W.$$

It follows that W is also a vector space over \mathbf{R} or \mathbf{C} , as appropriate, with respect to the restriction of the vector space operations on V to W .

If N is a norm on V , then it is easy to see that the restriction of N to W defines a norm on W . Of course, the metric on W associated to the restriction of N to W is the same as the restriction to W of the metric (12.2.3) associated to N on V .

If V is a Banach space with respect to N and W is a closed set in V with respect to (12.2.3), then W is a Banach space with respect to the restriction of N to W . More precisely, W is complete with respect to the restriction of (12.2.3) to W in this situation.

12.3 Hilbert spaces

Let V be a vector space over the real numbers. A real-valued function $\langle v, w \rangle$ defined for $v, w \in V$ is said to be an *inner product* on V if it satisfies the following three conditions. First, for each $w \in V$, $\langle v, w \rangle$ should be linear in V , as a mapping from V into \mathbf{R} . This means that

$$(12.3.1) \quad \langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle$$

for every $v, v' \in V$, and that

$$(12.3.2) \quad \langle t v, w \rangle = t \langle v, w \rangle$$

for every $v \in V$ and $t \in \mathbf{R}$.

Second, $\langle v, w \rangle$ should be symmetric in v and w , so that

$$(12.3.3) \quad \langle v, w \rangle = \langle w, v \rangle$$

for every $v, w \in V$. This implies that $\langle v, w \rangle$ is linear in w for every $v \in V$, because of the linearity in v .

Third,

$$(12.3.4) \quad \langle v, v \rangle > 0$$

for every $v \in V$ with $v \neq 0$. Of course, $\langle v, w \rangle = 0$ whenever $v = 0$ or $w = 0$, because of linearity in v and w .

12.3.1 The complex case

Similarly, if V is a vector space over the complex numbers, then a complex-valued function $\langle v, w \rangle$ defined for $v, w \in V$ is said to be an *inner product* if it satisfies the following three conditions. First, for every $w \in V$, $\langle v, w \rangle$ should be (complex) linear in v , as a mapping from V into \mathbf{C} . This means that (12.3.1) should hold for every $v, v' \in V$, as before, and that (12.3.2) should hold for every $v \in V$ and $t \in \mathbf{C}$.

Second, $\langle v, w \rangle$ should be Hermitian symmetric in v and w , which means that

$$(12.3.5) \quad \langle w, v \rangle = \overline{\langle v, w \rangle}$$

for every $v, w \in V$, where \bar{a} is the complex conjugate of a complex number a . Combining this with the first condition, we get that $\langle v, w \rangle$ is conjugate-linear in w for each $v \in V$, so that

$$(12.3.6) \quad \langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle$$

for every $w, w' \in V$, and

$$(12.3.7) \quad \langle v, tw \rangle = \bar{t} \langle v, w \rangle$$

for every $w \in V$ and $t \in \mathbf{C}$.

Observe that

$$(12.3.8) \quad \overline{\langle v, v \rangle} = \langle v, v \rangle$$

for every $v \in V$, which means that $\langle v, v \rangle \in \mathbf{R}$. The third condition is that (12.3.4) hold for every $v \in V$ with $v \neq 0$ again. As before, $\langle v, w \rangle = 0$ whenever $v = 0$ or $w = 0$.

12.3.2 Norms associated to inner products

Let $(V, \langle v, w \rangle)$ be a real or complex inner product space. Put

$$(12.3.9) \quad \|v\| = \langle v, v \rangle^{1/2}$$

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

$$(12.3.10) \quad |\langle v, w \rangle| \leq \|v\| \|w\|$$

for every $v, w \in V$, which is another version of the Cauchy–Schwarz inequality. More precisely, this can be shown using the fact that

$$(12.3.11) \quad \langle v + tw, v + tw \rangle = \|v + tw\|^2 \geq 0$$

for every $v, w \in V$ and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate.

Note that

$$(12.3.12) \quad \|tv\| = |t| \|v\|$$

for every $v \in V$ and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate. One can also check that

$$(12.3.13) \quad \|v + w\| \leq \|v\| + \|w\|$$

for every $v, w \in V$, using (12.3.10). Thus (12.3.9) defines a norm on V as a vector space over \mathbf{R} or \mathbf{C} , as appropriate.

If V is complete with respect to the metric associated to the norm (12.3.9), then V is said to be a *Hilbert space* with respect to the inner product $\langle v, w \rangle$. Otherwise, one can pass to a completion of V , as before.

12.4 Infinite series in Banach spaces

Let V be a vector space over the real or complex numbers, and let N be a norm on V . An infinite series

$$(12.4.1) \quad \sum_{j=1}^{\infty} v_j$$

with terms v_j in V for each $j \geq 1$ is said to *converge* in V with respect to N if the corresponding sequence of partial sums

$$(12.4.2) \quad \sum_{j=1}^l v_j$$

converges to an element of V with respect to the metric d_N associated to N . In this case, the value of the sum (12.4.1) is defined to be the limit of the sequence (12.4.2).

12.4.1 Some basic properties of infinite series in V

If (12.4.1) converges in V , and if $\sum_{j=1}^{\infty} w_j$ is another infinite series of elements of V that converges in V , then $\sum_{j=1}^{\infty} (v_j + w_j)$ converges in V too, with

$$(12.4.3) \quad \sum_{j=1}^{\infty} (v_j + w_j) = \sum_{j=1}^{\infty} v_j + \sum_{j=1}^{\infty} w_j.$$

This follows from the analogous statement for sums of convergent sequences in V , applied to the partial sums of these series.

Similarly, if (12.4.1) converges in V , and $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then $\sum_{j=1}^{\infty} t v_j$ converges in V as well, with

$$(12.4.4) \quad \sum_{j=1}^{\infty} t v_j = t \sum_{j=1}^{\infty} v_j.$$

This uses the fact that t times a convergent sequence in V converges to t times the limit of the sequence.

12.4.2 The Cauchy criterion

The condition that the sequence of partial sums (12.4.2) of an infinite series (12.4.1) be a Cauchy sequence with respect to the metric d_N associated to N is equivalent to saying that for each $\epsilon > 0$ there is a positive integer L such that

$$(12.4.5) \quad N\left(\sum_{j=k}^l v_j\right) < \epsilon$$

for every $l \geq k \geq L$. In particular, this holds when (12.4.1) converges in V , because a convergent sequence in any metric space is a Cauchy sequence.

Note that the Cauchy condition for the partial sums implies that

$$(12.4.6) \quad \lim_{l \rightarrow \infty} N(v_l) = 0,$$

by taking $k = l$ in (12.4.5). This is the same as saying that $\{v_j\}_{j=1}^{\infty}$ converges to 0 in V , with respect to d_N .

If V is a Banach space with respect to N , then the Cauchy condition for the partial sums implies that the series converges in V .

12.4.3 Absolute convergence with respect to the norm

An infinite series (12.4.1) with terms in V is said to be *absolutely convergent* with respect to N if

$$(12.4.7) \quad \sum_{j=1}^{\infty} N(v_j)$$

converges as an infinite series of nonnegative real numbers.

Observe that

$$(12.4.8) \quad N\left(\sum_{j=k}^l v_j\right) \leq \sum_{j=k}^l N(v_j)$$

for every $l \geq k \geq 1$, by the triangle inequality for N . If (12.4.1) converges absolutely with respect to N , then the sequence of partial sums (12.4.2) is a Cauchy sequence with respect to d_N . More precisely, one verifies (12.4.5) using (12.4.8) and the analogous Cauchy condition for the partial sums of (12.4.7). If V is a Banach space with respect to N , then it follows that (12.4.1) converges in V .

12.4.4 Orthogonal vectors in an inner product space

Now let $(V, \langle \cdot, \cdot \rangle)$ be a real or complex inner product space, and let us use the corresponding norm $\| \cdot \|$, as in the previous section. A pair v, w of vectors in V are said to be *orthogonal* if

$$(12.4.9) \quad \langle v, w \rangle = 0.$$

In this case, it is easy to see that

$$(12.4.10) \quad \|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Let (12.4.1) be an infinite series of pairwise-orthogonal vectors in V , so that v_j is orthogonal to v_k when $j \neq k$. This implies that

$$(12.4.11) \quad \left\| \sum_{j=k}^l v_j \right\|^2 = \sum_{j=k}^l \|v_j\|^2$$

for every $l \geq k \geq 1$.

If

$$(12.4.12) \quad \sum_{j=1}^{\infty} \|v_j\|^2$$

converges as an infinite series of nonnegative real numbers, then the sequence (12.4.2) is a Cauchy sequence in V . This uses (12.4.11) to get the Cauchy condition (12.4.5) from the analogous Cauchy condition for the partial sums of (12.4.12). If V is a Hilbert space with respect to $\langle \cdot, \cdot \rangle$, then it follows that (12.4.1) converges in V .

12.5 Bounded linear mappings

Let V and W be vector spaces, where more precisely V and W should both be defined over the real numbers, or both defined over the complex numbers. Also let N_V and N_W be norms on V and W , respectively. Thus

$$(12.5.1) \quad d_V(v, v') = N_V(v - v') \quad \text{and} \quad d_W(w, w') = N_W(w, w')$$

define metrics on V and W , respectively.

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

$$(12.5.2) \quad N_W(T(v)) \leq C N_V(v)$$

for every $v \in V$. In this case, we have that

$$(12.5.3) \quad \begin{aligned} d_W(T(u), T(v)) &= N_W(T(u) - T(v)) = N_W(T(u - v)) \\ &\leq C N_V(u - v) = C d_V(u, v) \end{aligned}$$

for every $u, v \in V$, so that T is Lipschitz with respect to the corresponding metrics on V and W .

12.5.1 Some related characterizations

Conversely, if a linear mapping T from V into W is Lipschitz with respect to the metrics $d_V(\cdot, \cdot)$ and $d_W(\cdot, \cdot)$, then it is easy to see that T is bounded with respect to N_V and N_W . More precisely, if a linear mapping T from V into W has the property that $N_W(T(v))$ is bounded on a ball in V of positive radius, then one can check that T is bounded as a linear mapping. This uses scalar multiplication to obtain (12.5.2) from the boundedness of $N_W(T(v))$ when $N_V(v)$ is less than a fixed radius.

In particular, if T is continuous at 0 with respect to the metrics $d_V(\cdot, \cdot)$ and $d_W(\cdot, \cdot)$ on V and W , then there is a $\delta > 0$ such that $N_W(T(v)) < 1$ when $v \in V$ satisfies $N_V(v) < \delta$. This implies that T is bounded as a linear mapping, as before.

If V is a finite-dimensional vector space, then every linear mapping T from V into W is bounded. Of course, this is trivial when $V = \{0\}$. Otherwise, there is a positive integer n such that V is isomorphic to \mathbf{R}^n or \mathbf{C}^n as a vector space over the real or complex numbers, as appropriate. This permits one to reduce to the same type of arguments as in Section 7.3.

12.5.2 The operator norm of a bounded linear mapping

Suppose that $V \neq \{0\}$, and that T is a bounded linear mapping from V into W with respect to N_V and N_W . The corresponding *operator norm* is defined by

$$(12.5.4) \quad \|T\|_{op} = \sup \left\{ \frac{N_W(T(v))}{N_V(v)} : v \in V, v \neq 0 \right\},$$

where the finiteness of the supremum follows from the boundedness of T . If $V = \{0\}$, then one can simply take $\|T\|_{op} = 0$.

Equivalently, $\|T\|_{op}$ is the same as the infimum of the nonnegative real numbers C such that (12.5.2) holds. As in Subsection 7.7.2, $\|T\|_{op}$ is also the same as $\text{Lip}(T)$, defined with respect to the metrics $d_V(\cdot, \cdot)$ and $d_W(\cdot, \cdot)$ associated to N_V and N_W .

If a is a real or complex number, as appropriate, then aT also defines a linear mapping from V into W , where $(aT)(v) = aT(v)$ for every $v \in V$. It is easy to see that aT is a bounded linear mapping when T is, with

$$(12.5.5) \quad \|aT\|_{op} = |a| \|T\|_{op}.$$

Note that $\|T\|_{op} = 0$ if and only if $T = 0$, which is to say that $T(v) = 0$ for every $v \in V$.

If T_1 and T_2 are linear mappings from V into W , then $T_1 + T_2$ defines a linear mapping from V into W as well, with

$$(12.5.6) \quad (T_1 + T_2)(v) = T_1(v) + T_2(v).$$

If T_1, T_2 are bounded as linear mappings with respect to N_V and N_W , then one can check that $T_1 + T_2$ is bounded as well, with

$$(12.5.7) \quad \|T_1 + T_2\|_{op} \leq \|T_1\|_{op} + \|T_2\|_{op}.$$

12.5.3 Spaces of bounded linear mappings

The space of linear mappings from V into W is a vector space over the real or complex numbers, as appropriate, with respect to pointwise addition and scalar multiplication of mappings. The space of bounded linear mappings from V into W is a linear subspace of the space of all linear mappings from V into W , as in the preceding paragraph. The operator norm (12.5.4) defines a norm on the space of bounded linear mappings from V into W , as a vector space over the real or complex numbers, as appropriate.

12.5.4 Compositions of bounded linear mappings

Let Z be another vector space, which is defined over the real or complex numbers, depending on whether V and W are defined over the real or complex numbers. Also let N_Z be a norm on Z , let T_1 be a bounded linear mapping from V into W , and let T_2 be a bounded linear mapping from W into Z . Under these conditions, one can verify that the composition $T_2 \circ T_1$ is bounded as a linear mapping from V into Z , with

$$(12.5.8) \quad \|T_2 \circ T_1\|_{op} \leq \|T_1\|_{op} \|T_2\|_{op}.$$

Here the various operator norms are defined with respect to the given norms on V , W , and Z , as appropriate.

12.6 Dual spaces

Let V be a vector space over the real or complex numbers. A *linear functional* on V is a linear mapping from V into \mathbf{R} or \mathbf{C} , as appropriate, considered as a one-dimensional vector space over itself. The space of all linear functionals on V may be denoted V^{alg} , and is called the *algebraic dual* of V . This is a linear subspace of the space of all real or complex-valued functions on V , as a vector space over \mathbf{R} or \mathbf{C} , as appropriate.

12.6.1 Linear functionals on $c_{00}(X, \mathbf{R})$, $c_{00}(X, \mathbf{C})$

Let X be a nonempty set, and suppose for the moment that V is the space $c_{00}(X, \mathbf{R})$ or $c_{00}(X, \mathbf{C})$ of real or complex-valued functions on X with finite support, as in Section 1.6, considered as a vector space over \mathbf{R} or \mathbf{C} , as appropriate. If f and g are real or complex-valued functions on X , as appropriate, and f has finite support in X , then it is easy to see that fg has finite support in X too. In this case, we put

$$(12.6.1) \quad \lambda_g(f) = \sum_{x \in X} f(x) g(x),$$

where the sum over X on the right is defined by reducing to a finite sum, as in Subsection 1.6.1. This defines λ_g as a linear functional on V .

If $x \in X$, then let δ_x be the real-valued function defined on X by putting δ_x equal to 1 at x , and equal to 0 at every other element of X , as in Subsection 1.10.1. Observe that

$$(12.6.2) \quad \lambda_g(\delta_x) = g(x)$$

for every $x \in X$. If λ is any linear functional on V , then

$$(12.6.3) \quad \lambda = \lambda_g$$

for a unique real or complex-valued function g on X , as appropriate. More precisely, one can take

$$(12.6.4) \quad g(x) = \lambda(\delta_x)$$

for each $x \in X$.

12.6.2 Bounded linear functionals

Let V be a vector space over the real or complex numbers with a norm N_V . A linear functional λ on V is said to be *bounded* with respect to N_V if λ is bounded as a linear mapping from V into \mathbf{R} or \mathbf{C} , as appropriate, the previous section. This uses the usual absolute value function on \mathbf{R} or \mathbf{C} as the norm on \mathbf{R} or \mathbf{C} , as appropriate, considered as a one-dimensional vector space over itself.

The space of bounded linear functionals on V may be denoted V' or V^* . This is a linear subspace of the space V^{alg} of all linear functionals on V , as in Subsection 12.5.3. If λ is a bounded linear functional on V , then the corresponding operator norm of λ may be denoted $N_{V'}(\lambda)$ or $N_{V^*}(\lambda)$. This is the *dual norm* associated to N_V .

If V has finite dimension, then every linear functional on V is bounded with respect to N_V , as in Subsection 12.5.1.

12.6.3 Using an inner product

Suppose for the moment again that $\langle \cdot, \cdot \rangle$ is an inner product on V , and that N_V is the norm associated to the inner product, as in Subsection 12.3.2. If $w \in V$, then

$$(12.6.5) \quad \mu_w(v) = \langle v, w \rangle$$

defines a linear functional on V . More precisely, this is a bounded linear functional on V , because of the Cauchy-Schwarz inequality, as in Subsection 12.3.2. It is not too difficult to show that the dual norm of μ_w is equal to the norm of w . This uses the fact that

$$(12.6.6) \quad \mu_w(w) = \langle w, w \rangle$$

is the same as the norm of w squared.

If V is a Hilbert space, then a famous theorem of Riesz states that every bounded linear functional on V is of the form (12.6.5) for a unique $w \in V$. A version of this will be mentioned in the next section.

12.6.4 A part of the Hahn–Banach theorem

Let V be any real or complex vector space with a norm N_V again. If $v \in V$, then part of the *Hahn–Banach theorem* implies that there is a bounded linear functional λ on V with respect to N_V such that

$$(12.6.7) \quad \lambda(v) = N_V(v)$$

and the dual norm of λ is less than or equal to one. If $v \neq 0$, then it follows that the dual norm of λ is equal to one.

If N_V is the norm associated to an inner product on V , then this may be obtained from the remarks in the preceding subsection.

12.7 Some bounded linear functionals

Let X be a nonempty set, and let g be a bounded real or complex-valued function on X . If f is a summable real or complex-valued function on X , as in Section 11.6, then one can check that the product of f and g is summable on X as well. In fact,

$$(12.7.1) \quad \|fg\|_1 \leq \|f\|_1 \|g\|_\infty.$$

Using this, $\lambda_g(f)$ may be defined as a real or complex number, as appropriate, as in (12.6.1). This uses the definition of the sum of a summable real or complex-valued function on X as in Sections 11.7 and 11.8. If X is the set \mathbf{Z}_+ of positive integers, then the sum may be considered as an absolutely convergent infinite series. Of course, if X has only finitely many elements, then the sum over X may be defined more directly.

One can check that λ_g defines a bounded linear functional on $\ell^1(X, \mathbf{R})$ or $\ell^1(X, \mathbf{C})$, as appropriate, with respect to the usual ℓ^1 norm. The corresponding dual norm of λ_g is less than or equal to $\|g\|_\infty$, because of (12.7.1). One can verify that the dual norm of λ_g is equal to $\|g\|_\infty$.

It is well known that every bounded linear functional on $\ell^1(X, \mathbf{R})$ or $\ell^1(X, \mathbf{C})$ corresponds to a unique bounded real or complex-valued function g on X in this way, as appropriate.

12.7.1 Some more examples

Now let g be a summable real or complex-valued function on X . If f is a bounded real or complex-valued function on X , then fg is summable on X , with

$$(12.7.2) \quad \|fg\|_1 \leq \|f\|_\infty \|g\|_1.$$

One can use this to define $\lambda_g(f)$ as a real or complex number, as appropriate, as in (12.6.1), as before.

More precisely, λ_g is a bounded linear functional on $\ell^\infty(X, \mathbf{R})$ or $\ell^\infty(X, \mathbf{C})$, as appropriate, with respect to the supremum norm. The corresponding dual

norm of λ_g is less than or equal to $\|g\|_1$, because of (12.7.2). It is not too difficult to show that the dual norm of λ_g is equal to $\|g\|_1$.

One can restrict λ_g to the space of real-or complex-valued functions on X that vanish at infinity, as in Section 2.5. This defines a bounded linear functional on $c_0(X, \mathbf{R})$ or $c_0(X, \mathbf{C})$, as appropriate, with respect to the supremum norm. It is not too difficult to show that the dual norm of the restriction of λ_g to $c_0(X, \mathbf{R})$ or $c_0(X, \mathbf{C})$, as appropriate, is also equal to $\|g\|_1$.

It is well known that every bounded linear functional on $c_0(X, \mathbf{R})$ or $c_0(X, \mathbf{C})$ with respect to the supremum norm corresponds to a unique summable real or complex-valued function g on X in this way, as appropriate.

The spaces $\ell^2(X, \mathbf{R})$, $\ell^2(X, \mathbf{C})$ are Hilbert spaces with respect to the inner products defined in Section 11.20. The Riesz representation theorem can be shown more directly in this case.

Appendix A

Some more on mappings, metrics, and norms

A.1 A nice inequality

Let a be a positive real number, with $a \leq 1$. If r and t are nonnegative real numbers, then it is well known that

$$(A.1.1) \quad (r+t)^a \leq r^a + t^a.$$

To see this, observe first that

$$(A.1.2) \quad \max(r, t) \leq (r^a + t^a)^{1/a}.$$

Using this, we get that

$$(A.1.3) \quad r+t \leq \max(r, t)^{1-a} (r^a + t^a) \leq (r^a + t^a)^{((1-a)/a)+1} = (r^a + t^a)^{1/a}.$$

This is equivalent to (A.1.1).

A.1.1 Snowflake transforms

Let $(X, d(x, y))$ be a metric space. If $0 < a \leq 1$, then it is easy to see that

$$(A.1.4) \quad d(x, y)^a$$

is a metric on X as well, using (A.1.1). This was mentioned in Subsection 1.1.2 when $a = 1/2$.

Let us use $B_d(x, r)$, $\overline{B}_d(x, r)$ to denote the open and closed balls in X centered at $x \in X$ with radius r with respect to $d(\cdot, \cdot)$, respectively, as in Section 1.9, and $B_{d^a}(x, r)$, $\overline{B}_{d^a}(x, r)$ for the open and closed balls in X centered at x with radius r with respect to $d(\cdot, \cdot)^a$, respectively. Observe that

$$(A.1.5) \quad B_{d^a}(x, r^a) = B_d(x, r)$$

for every $r > 0$, and that

$$(A.1.6) \quad \overline{B}_{d^a}(x, r^a) = \overline{B}_d(x, r)$$

for every $r \geq 0$.

It is easy to see that the identity mapping on X is uniformly continuous as a mapping from X equipped with $d(\cdot, \cdot)$ into X equipped with $d(\cdot, \cdot)^a$. Similarly, the identity mapping on X is uniformly continuous as a mapping from X equipped with $d(\cdot, \cdot)^a$ into X equipped with $d(\cdot, \cdot)$. These statements were also mentioned in Subsection 1.1.2 when $a = 1/2$.

If $a > 1$, then one can verify that

$$(A.1.7) \quad |x - y|^a$$

is not a metric on the real line. This corresponds to the first part of Exercise 11 at the end of Chapter 2 in [210] when $a = 2$.

A.2 Some more Lipschitz conditions

Let (X, d_X) and (Y, d_Y) be metric spaces, and let α be a positive real number. A mapping f from X into Y is said to be *Lipschitz of order α* if there is a nonnegative real number C such that

$$(A.2.1) \quad d_Y(f(x), f(w)) \leq C d_X(x, w)^\alpha$$

for every $x, w \in X$. This is the same as a Lipschitz mapping as in Section 1.1 when $\alpha = 1$. One may also say that f is *Hölder continuous of order α* in this case.

Lipschitz mappings of any order are uniformly continuous, as before. Of course, (A.2.1) holds with $C = 0$ if and only if f is constant on X .

Let a be a positive real number, and suppose that

$$(A.2.2) \quad d_X(x, w)^a \text{ is a metric on } X.$$

This holds automatically when $a \leq 1$, as in the previous section. Note that (A.2.1) is the same as saying that

$$(A.2.3) \quad d_Y(f(x), f(w)) \leq C (d_X(x, w)^a)^{\alpha/a}$$

for every $x, w \in X$. This means that

$$(A.2.4) \quad f \text{ is Lipschitz of order } \alpha \text{ with respect to } d_X \text{ on } X$$

if and only if

$$(A.2.5) \quad f \text{ is Lipschitz of order } \alpha/a \text{ with respect to } d_X(\cdot, \cdot)^a \text{ on } X,$$

and with the same constant C .

Similarly, let b be a positive real number, and suppose that

$$(A.2.6) \quad d_Y(\cdot, \cdot)^b \text{ is a metric on } Y.$$

Clearly (A.2.1) is the same as saying that

$$(A.2.7) \quad d_Y(f(x), f(w))^b \leq C^b d_X(x, w)^{\alpha b}$$

for every $x, w \in X$. Thus

$$(A.2.8) \quad f \text{ is Lipschitz of order } \alpha \text{ with constant } C \text{ with respect to } d_Y \text{ on } Y$$

if and only if

$$(A.2.9) \quad \begin{aligned} f \text{ is Lipschitz of order } \alpha b \text{ with constant } C^b \\ \text{with respect to } d_Y(\cdot, \cdot)^b \text{ on } Y. \end{aligned}$$

A.2.1 Lipschitz functions of order $\alpha > 1$ on \mathbf{R}

Suppose that $X = Y = \mathbf{R}$, with the standard Euclidean metric. If f is Lipschitz of order $\alpha > 1$, then

$$(A.2.10) \quad f \text{ is constant on } \mathbf{R}.$$

More precisely, it is easy to see that the derivative of f is equal to 0 at every point in \mathbf{R} under these conditions.

A.3 Another nice inequality

Let X be a nonempty set, and let f be a real or complex-valued function on X with finite support, as in Section 1.6. If p is a positive real number, then put

$$(A.3.1) \quad \|f\|_p = \left(\sum_{x \in X} |f(x)|^p \right)^{1/p}.$$

This is the same as in Subsection 1.6.1 when $p = 1$ or 2 . If $X = \{1, \dots, n\}$ for some positive integer n , then this corresponds to an analogous expression in Subsection 1.4.1. We also put

$$(A.3.2) \quad \|f\|_\infty = \max_{x \in X} |f(x)|,$$

as before.

Observe that

$$(A.3.3) \quad \|f\|_\infty \leq \|f\|_p.$$

If $0 < p_1 \leq p_2 < \infty$, then we would like to check that

$$(A.3.4) \quad \|f\|_{p_2} \leq \|f\|_{p_1}.$$

Clearly

$$(A.3.5) \quad \|f\|_{p_2}^{p_2} = \sum_{x \in X} |f(x)|^{p_2} \leq \|f\|_{\infty}^{p_2 - p_1} \sum_{x \in X} |f(x)|^{p_1} = \|f\|_{\infty}^{p_2 - p_1} \|f\|_{p_1}^{p_1}.$$

This implies that

$$(A.3.6) \quad \|f\|_{p_2}^{p_2} \leq \|f\|_{p_1}^{p_2},$$

because of (A.3.3). It follows that (A.3.4) holds, as desired.

It is easy to see that (A.1.1) follows from (A.3.4), with $p_1 = a$, $p_2 = 1$, and where X has two elements.

A.3.1 The limit of $\|f\|_p$ as $p \rightarrow \infty$

Observe that

$$(A.3.7) \quad \|f\|_p \leq (\# \text{supp } f)^{1/p} \|f\|_{\infty},$$

where $\# \text{supp } f$ is the number of elements in the support of f . One can use this and (A.3.3) to get that

$$(A.3.8) \quad \|f\|_p \rightarrow \|f\|_{\infty} \text{ as } p \rightarrow \infty.$$

A.3.2 A related metric for $p \leq 1$

Of course, $\|f\|_p = 0$ if and only if $f = 0$ on X , and

$$(A.3.9) \quad \|t f\|_p = |t| \|f\|_p$$

for every $t \in \mathbf{R}$ or \mathbf{C} , as appropriate.

If g is another real or complex-valued function on X with finite support and $0 < p \leq 1$, then

$$\begin{aligned} \|f + g\|_p^p &= \sum_{x \in X} |f(x) + g(x)|^p \leq \sum_{x \in X} (|f(x)| + |g(x)|)^p \\ (A.3.10) \quad &\leq \sum_{x \in X} (|f(x)|^p + |g(x)|^p) = \sum_{x \in X} |f(x)|^p + \sum_{x \in X} |g(x)|^p \\ &= \|f\|_p^p + \|g\|_p^p, \end{aligned}$$

using (A.1.1) in the third step, with $a = p$. This implies that

$$(A.3.11) \quad \|f - g\|_p^p$$

defines a metric on $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$ when $0 < p \leq 1$.

A.4 Some functions on $[0, 1]$

Let j be a positive integer, and let f_j be the real-valued function defined on $[0, 1]$ by

$$(A.4.1) \quad \begin{aligned} f_j(x) &= 2jx && \text{when } 0 \leq x \leq 1/(2j) \\ &= 2 - 2jx && \text{when } 1/(2j) \leq x \leq 1/j \\ &= 0 && \text{when } 1/j \leq x \leq 1. \end{aligned}$$

Thus

$$(A.4.2) \quad f_j(0) = 0, \quad f_j(1/(2j)) = 1, \quad \text{and} \quad f_j(1/j) = 0,$$

with the overlapping definitions of $f_j(x)$ agreeing at $x = 1/(2j), 1/j$. Equivalently, f_j is defined to be linear on $[0, 1/(2j)]$ and $[1/(2j), 1/j]$, with these values at the endpoints.

Note that f_j is continuous on $[0, 1]$. One can check that

$$(A.4.3) \quad \{f_j\}_{j=1}^{\infty} \text{ converges to 0 pointwise on } [0, 1],$$

and not uniformly. In fact,

$$(A.4.4) \quad \|f_j\|_{\infty} = 1$$

for every j , where $\|\cdot\|_{\infty}$ is the supremum norm on $C([0, 1], \mathbf{R})$, as before.

Observe that

$$(A.4.5) \quad \|f_j\|_1 = \int_0^1 f_j(x) dx = 2 \int_0^{1/(2j)} f_j(x) dx = 2j(1/(2j))^2 = 1/(2j)$$

for every j , where $\|\cdot\|_1$ is as in (1.15.4). More precisely,

$$(A.4.6) \quad f_j(x) = f_j((1/j) - x) \quad \text{when } 0 \leq x \leq 1/j,$$

which is to say that $f_j(x)$ is symmetric about $x = 1/(2j)$ on $[0, 1/j]$. This implies that the integrals of $f_j(x)$ over $[0, 1/(2j)]$ and $[1/(2j), 1/j]$ are the same. Of course, (A.4.5) is the same as the area of the triangle determined by the graph of f_j on $[0, 1/j]$.

Similarly,

$$(A.4.7) \quad \begin{aligned} \|f_j\|_2^2 &= \int_0^1 f_j(x)^2 dx = 2 \int_0^{1/(2j)} f_j(x)^2 dx \\ &= (2/3)(2j)^2(1/(2j))^3 = 1/(3j) \end{aligned}$$

for every j , where $\|\cdot\|_2$ is as in (1.15.6). This uses (A.4.6) in the second step, to get that the integrals of $f_j(x)^2$ over $[0, 1/(2j)]$ and $[1/(2j), 1/j]$ are the same. It follows that

$$(A.4.8) \quad \|f_j\|_2 = 1/(3j)^{1/2}$$

for each j .

A.4.1 Some related sequences of functions

If $\alpha \in \mathbf{R}$, then one can verify that

$$(A.4.9) \quad \{j^{-\alpha} f_j\}_{j=1}^{\infty} \text{ converges to 0 pointwise on } [0, 1].$$

Note that

$$(A.4.10) \quad \|j^{-\alpha} f_j\|_{\infty} = j^{-\alpha} \|f_j\|_{\infty} = j^{-\alpha}$$

for every j , by (A.4.4). This implies that $\{j^{-\alpha} f_j\}_{j=1}^{\infty}$ is bounded with respect to the supremum metric exactly when $\alpha \geq 0$, and that $\{f_j\}_{j=1}^{\infty}$ converges to 0 with respect to the supremum metric exactly when $\alpha > 0$.

Similarly,

$$(A.4.11) \quad \|j^{-\alpha} f_j\|_1 = j^{-\alpha} \|f_j\|_1 = (1/2) j^{-1-\alpha}$$

for every j , by (A.4.5). This means that $\{j^{-\alpha} f_j\}_{j=1}^{\infty}$ is bounded with respect to the metric d_1 on $C([0, 1], \mathbf{R})$ associated to $\|\cdot\|_1$ as in (1.15.10) exactly when $\alpha \geq -1$, and that $\{j^{-\alpha} f_j\}_{j=1}^{\infty}$ converges to 0 with respect to d_1 exactly when $\alpha > -1$.

We also have that

$$(A.4.12) \quad \|j^{-\alpha} f_j\|_2 = j^{-\alpha} \|f_j\|_2 = (1/\sqrt{3}) j^{-(1/2)-\alpha}$$

for every j , by (A.4.8). It follows that $\{j^{-\alpha} f_j\}_{j=1}^{\infty}$ is bounded with respect to the metric d_2 on $C([0, 1], \mathbf{R})$ associated to $\|\cdot\|_2$ as in (1.15.11) if and only if $\alpha \geq -1/2$, and that $\{j^{-\alpha} f_j\}_{j=1}^{\infty}$ converges to 0 with respect to d_2 if and only if $\alpha > -1/2$.

A.5 Some Cauchy sequences

Let j be a positive integer, and let f_j be the real-valued function defined on $[0, 1]$ by

$$(A.5.1) \quad \begin{aligned} f_j(x) &= 0 && \text{when } 0 \leq x \leq 1/2 - 1/(2j) \\ &= 2j(x - (1/2 - 1/(2j))) && \text{when } 1/2 - 1/(2j) \leq x \leq 1/2 \\ &= 1 && \text{when } 1/2 \leq x \leq 1. \end{aligned}$$

The second case says that f_j is linear in that range, with the same values at the endpoints as in the other two cases. In particular, f_j is continuous on $[0, 1]$ for each j .

It is easy to see that $\{f_j\}_{j=1}^{\infty}$ converges pointwise on $[0, 1]$ to the real-valued function f defined on $[0, 1]$ by

$$(A.5.2) \quad \begin{aligned} f(x) &= 0 && \text{when } 0 \leq x < 1/2 \\ &= 1 && \text{when } 1/2 \leq x \leq 1. \end{aligned}$$

Note that $\{f_j\}_{j=1}^{\infty}$ does not converge to f uniformly on $[0, 1]$.

However, one can check that $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence in $C([0, 1], \mathbf{R})$ with respect to the metric d_1 defined in (1.15.10). Basically, $\{f_j\}_{j=1}^\infty$ converges to f with respect to a metric like d_1 on a larger space. Similarly, one can verify that $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence in $C([0, 1], \mathbf{R})$ with respect to the metric d_2 defined in (1.15.11).

A.5.1 Another sequence of functions

Let j be a positive integer again, and let f_j be the real-valued function defined on $[0, 1]$ by

$$(A.5.3) \quad f_j(x) = \min(1/\sqrt{x}, j),$$

where the right side is interpreted as being equal to j when $x = 0$. Note that f_j is continuous on $[0, 1]$ for each j . One can check that $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence in $C([0, 1], \mathbf{R})$ with respect to d_1 . However, $\{f_j\}_{j=1}^\infty$ is not bounded as a sequence in $C([0, 1], \mathbf{R})$ with respect to d_2 .

A.5.2 Some functions on \mathbf{Z}_+

Now let f be a real or complex-valued function on $X = \mathbf{Z}_+$. If j is a positive integer, then let f_j be the real or complex-valued function, as appropriate, defined on \mathbf{Z}_+ by

$$(A.5.4) \quad \begin{aligned} f_j(l) &= f(l) && \text{when } l \leq j \\ &= 0 && \text{when } l > j. \end{aligned}$$

Clearly $\{f_j\}_{j=1}^\infty$ converges to f pointwise on \mathbf{Z}_+ . By construction, the support of f_j has only finitely many elements for each j .

If

$$(A.5.5) \quad \lim_{l \rightarrow \infty} f(l) = 0,$$

then $\{f_j\}_{j=1}^\infty$ converges to f uniformly on \mathbf{Z}_+ . Note that $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence in $c_{00}(\mathbf{Z}_+, \mathbf{R})$ or $c_{00}(\mathbf{Z}_+, \mathbf{C})$, as appropriate, with respect to the metric $d_\infty(\cdot, \cdot)$ defined in Subsection 1.6.1 under these conditions.

Suppose for the moment that

$$(A.5.6) \quad \sum_{l=1}^{\infty} |f(l)|$$

converges as an infinite series of nonnegative real numbers. In this case, one can check that $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence in $c_{00}(\mathbf{Z}_+, \mathbf{R})$ or $c_{00}(\mathbf{Z}_+, \mathbf{C})$, as appropriate, with respect to the metric $d_1(\cdot, \cdot)$ defined in Subsection 1.6.1.

Similarly, suppose instead that

$$(A.5.7) \quad \sum_{l=1}^{\infty} |f(l)|^2$$

converges as an infinite series of nonnegative real numbers. One can verify that $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence in $c_{00}(\mathbf{Z}_+, \mathbf{R})$ or $c_{00}(\mathbf{Z}_+, \mathbf{C})$, as appropriate, with respect to the metric $d_2(\cdot, \cdot)$ defined in Subsection 1.6.1 in this situation.

A.6 Norms and convexity

If the reader is familiar with the abstract notion of a vector space, then let V be a vector space over the real numbers. Otherwise, one can consider particular situations like $V = \mathbf{R}^n$ for some positive integer n , or a vector space of real-valued functions on a nonempty set.

As usual, a subset E of V is said to be *convex* if for every $v, w \in E$ and $t \in \mathbf{R}$ with $0 \leq t \leq 1$ we have that

$$(A.6.1) \quad tv + (1 - t)w \in E.$$

A.6.1 The definition of a seminorm

Let N be a nonnegative real-valued function on V such that

$$(A.6.2) \quad N(tv) = |t|N(v)$$

for every $t \in \mathbf{R}$ and $v \in V$. If N also satisfies the triangle inequality

$$(A.6.3) \quad N(v + w) \leq N(v) + N(w)$$

for every $v, w \in V$, then N is said to be a *seminorm* on V . Thus a norm on V is the same as a seminorm N such that $N(v) > 0$ for every $v \in V$ with $v \neq 0$.

Put

$$(A.6.4) \quad \overline{B}_N = \{v \in V : N(v) \leq 1\},$$

which is the closed unit ball in V with respect to N . If N is a seminorm on V , then it is easy to see that \overline{B}_N is a convex subset of V . Conversely, if N is a nonnegative real-valued function on V that satisfies (A.6.2), and if \overline{B}_N is a convex set in V , then N satisfies (A.6.3), and hence is a seminorm on V . This is not too difficult to show, directly from the definitions.

A.6.2 Some norms on $c_{00}(X, \mathbf{R})$

Let X be a nonempty set, and let $c_{00}(X, \mathbf{R})$ be the space of real-valued functions on X with finite support, as in Section 1.6. If $f \in c_{00}(X, \mathbf{R})$ and p is a positive real number, then $\|f\|_p$ may be defined as in Section A.3. This satisfies the first two conditions in the definition of a norm on $c_{00}(X, \mathbf{R})$, as before.

If $p \geq 1$, then one can show that the corresponding closed unit ball

$$(A.6.5) \quad \overline{B}_p = \{f \in c_{00}(X, \mathbf{R}) : \|f\|_p \leq 1\}$$

is a convex set in $c_{00}(X, \mathbf{R})$. More precisely, (A.6.5) is the same as taking

$$(A.6.6) \quad \overline{B}_p = \left\{ f \in c_{00}(X, \mathbf{R}) : \|f\|_p^p = \sum_{x \in X} |f(x)|^p \leq 1 \right\}.$$

One can show that this is a convex subset of $c_{00}(X, \mathbf{R})$ when $p \geq 1$, using the convexity of r^p for $r \geq 0$ when $p \geq 1$. This implies that $\|f\|_p$ defines a norm on $c_{00}(X, \mathbf{R})$ when $p \geq 1$, as in the preceding subsection. If $0 < p < 1$, then (A.6.5) is not convex in $c_{00}(X, \mathbf{R})$ when X has at least two elements.

A.6.3 Complex vector spaces

A vector space over the complex numbers may be considered as a vector space over the real numbers as well, to which the earlier remarks about convexity can be applied. In particular, there are analogues of the statements in the previous subsection for the space $c_{00}(X, \mathbf{C})$ of complex-valued functions on X with finite support.

A.6.4 More on seminorms

Let n be a positive integer, and let N be a seminorm on \mathbf{R}^n . The closed unit ball \overline{B}_N in \mathbf{R}^n with respect to N is convex, as before, as is the open unit ball

$$(A.6.7) \quad B_N = \{v \in \mathbf{R}^n : N(v) < 1\}.$$

Note that B_N and \overline{B}_N are also symmetric about the origin, which is to say that they are invariant under the mapping $v \mapsto -v$ on \mathbf{R}^n .

One can show that N is continuous as a real-valued function on \mathbf{R}^n , and in fact it is Lipschitz with respect to the standard Euclidean metric on \mathbf{R}^n , as in Subsection 6.11.1. This implies that B_N is an open set, and \overline{B}_N is a closed set, with respect to the standard Euclidean metric on \mathbf{R}^n .

If N is a norm on \mathbf{R}^n , then the standard Euclidean norm on \mathbf{R}^n is bounded by a constant multiple of N , as in Subsection 6.11.3. It follows that B_N and \overline{B}_N are bounded sets with respect to the standard Euclidean metric on \mathbf{R}^n in this case.

A.7 Path-connected sets

Let (X, d_X) be a metric space, and let a, b be real numbers with $a < b$. Suppose that

$$(A.7.1) \quad p \text{ is a continuous mapping from the closed interval } [a, b] \text{ into } X,$$

with respect to the restriction of the standard Euclidean metric on \mathbf{R} to $[a, b]$. Under these conditions, one can check that

$$(A.7.2) \quad p([a, b]) \text{ is a connected subset of } X.$$

More precisely, it is well known that $[a, b]$ is a connected subset of the real line, with respect to the standard metric. This implies that $[a, b]$ is connected as a subset of itself, with respect to the restriction of the standard metric on \mathbf{R} to $[a, b]$, as in Subsection 8.12.1. It follows that (A.7.2) holds, because of the well-known theorem that continuous mappings send connected sets to connected sets. Alternatively, one can extend p to a continuous mapping from \mathbf{R} into X , by putting $p(t) = p(a)$ when $t < a$, and $p(t) = p(b)$ when $t > b$.

A.7.1 The definition of path connectedness

A subset E of X is said to be *path connected* if for every pair of points $x, w \in E$ there are real numbers a, b with $a < b$ and a continuous mapping p from $[a, b]$ into X such that $p(a) = x$, $p(b) = w$, and

$$(A.7.3) \quad p([a, b]) \subseteq E.$$

It is well known and not too difficult to show that

$$(A.7.4) \quad \text{path-connected sets are connected,}$$

using (A.7.2).

If E is a convex set in \mathbf{R}^n for some positive integer n , then

$$(A.7.5) \quad E \text{ is path connected,}$$

with respect to the standard Euclidean metric on \mathbf{R}^n . Note that connected subsets of \mathbf{R} are convex.

A.7.2 Connected open sets in \mathbf{R}^n

If U is a connected open set in \mathbf{R}^n , then it is well known that

$$(A.7.6) \quad U \text{ is path connected.}$$

To see this, let $x \in U$ be given, and let U_x be the set of $w \in U$ for which there is a continuous path in U from x to w , as before. One can check that

$$(A.7.7) \quad U_x \text{ is an open set in } \mathbf{R}^n.$$

Similarly, one can verify that

$$(A.7.8) \quad U \setminus U_x \text{ is an open set in } \mathbf{R}^n.$$

If U is connected, then it follows that

$$(A.7.9) \quad U_x = U.$$

It is well known that there are connected subsets of \mathbf{R}^2 that are not path connected.

A.7.3 Closures of connected sets

If E is a connected set in any metric space X , then it is well known and not too difficult to show that

$$(A.7.10) \quad \overline{E} \text{ is connected in } X$$

too. However, there are path-connected subsets of \mathbf{R}^2 whose closures are not path connected.

Let (Y, d_Y) be another metric space, and let f be a continuous mapping from X into Y . If E is a path-connected subset of X , then it is easy to see that

$$(A.7.11) \quad f(E) \text{ is path connected in } Y.$$

A.8 Bounded vector-valued functions

Let X be a nonempty set, and let m be a positive integer. Consider the spaces

$$(A.8.1) \quad c(X, \mathbf{R}^m), \quad c(X, \mathbf{C}^m)$$

of functions on X with values in \mathbf{R}^m , \mathbf{C}^m , respectively. These spaces with $m = 1$ were discussed in Section 3.10.

As before, if f and g are functions on X with values in \mathbf{R}^m or \mathbf{C}^m , then $f + g$ defines a function on X with values in \mathbf{R}^m or \mathbf{C}^m , as appropriate. Similarly, if t is a real or complex number, as appropriate, then tf is a function on X with values in \mathbf{R}^m or \mathbf{C}^m , as appropriate. As usual, $c(X, \mathbf{R}^m)$ and $c(X, \mathbf{C}^m)$ are basic classes of examples of vector spaces over the real and complex numbers, respectively. We shall also be interested in linear subspaces of these spaces, which are subsets of the spaces that contain 0 and are closed under addition and scalar multiplication, as before.

If f is a function on X with values in \mathbf{R}^m or \mathbf{C}^m , then let $f_j(x)$ be the j th coordinate of $f(x)$ for each $j = 1, \dots, m$ and $x \in X$. Thus f_j is a real or complex-valued function on X , as appropriate, for each j . Of course, any m -tuple of real or complex-valued functions on X determines a function on X with values in \mathbf{R}^m or \mathbf{C}^m , as appropriate, in this way.

A.8.1 Using a norm N on \mathbf{R}^m or \mathbf{C}^m

Let N be a norm on \mathbf{R}^m or \mathbf{C}^m , so that $d_N(v, w) = N(v - w)$ defines a metric on \mathbf{R}^m or \mathbf{C}^m , as appropriate. Using this, we get the corresponding space

$$(A.8.2) \quad \mathcal{B}(X, \mathbf{R}^m) = \mathcal{B}_N(X, \mathbf{R}^m) \text{ or } \mathcal{B}(X, \mathbf{C}^m) = \mathcal{B}_N(X, \mathbf{C}^m)$$

of mappings from X into \mathbf{R}^m or \mathbf{C}^m , as appropriate, that are bounded with respect to $d_N(\cdot, \cdot)$, as in Section 1.11. This space does not depend on the particular norm N , because any norm on \mathbf{R}^m or \mathbf{C}^m can be compared with the standard Euclidean norm up to constant factors, as in Subsections 6.10.3 and 6.11.3. The corresponding supremum metric does depend on N , but it can be compared with the supremum metric associated to the standard Euclidean norm on \mathbf{R}^m or \mathbf{C}^m , as appropriate, with the same constant factors.

We may also use the notation

$$(A.8.3) \quad \ell^\infty(X, \mathbf{R}^m) = \ell_N^\infty(X, \mathbf{R}^m) \text{ or } \ell^\infty(X, \mathbf{C}^m) = \ell_N^\infty(X, \mathbf{C}^m)$$

for (A.8.2), as appropriate. Note that a function f on X with values in \mathbf{R}^m or \mathbf{C}^m , as appropriate, is bounded with respect to $d_N(\cdot, \cdot)$ if and only if

$$(A.8.4) \quad N(f(x))$$

is bounded as a real-valued function on X . This happens if and only if f_j is bounded as a real or complex-valued function on X , as appropriate, for each j , because of the usual comparisons of N with the standard Euclidean norm. It is easy to see that (A.8.3) is a linear subspace of $c(X, \mathbf{R}^m)$ or $c(X, \mathbf{C}^m)$, as appropriate.

A.8.2 Supremum norms associated to N

If f is an element of (A.8.3), as appropriate, then put

$$(A.8.5) \quad \|f\|_{\infty, N} = \sup\{N(f(x)) : x \in X\}.$$

This is a nonnegative real number, which is equal to 0 exactly when $f = 0$ on X . One can check that

$$(A.8.6) \quad \|tf\|_{\infty, N} = |t| \|f\|_{\infty, N}$$

for every $t \in \mathbf{R}$ or \mathbf{C} , as appropriate. If g is another element of (A.8.3), as appropriate, then one can verify that

$$(A.8.7) \quad \|f + g\|_{\infty, N} \leq \|f\|_{\infty, N} + \|g\|_{\infty, N}.$$

This means that (A.8.5) defines a norm on (A.8.3), as appropriate. This is the *supremum norm* associated to N . This can be compared with the supremum norm associated to the standard Euclidean norm on \mathbf{R}^m or \mathbf{C}^m , as appropriate, with the same constant factors as before. Note that

$$(A.8.8) \quad \|f - g\|_{\infty, N} = \sup\{N(f(x) - g(x)) : x \in X\}$$

is the same as the supremum metric on (A.8.3), as appropriate, corresponding to $d_N(\cdot, \cdot)$.

A.8.3 Continuous vector-valued functions

Suppose now that (X, d_X) is a nonempty metric space, and consider the space

$$(A.8.9) \quad C(X, \mathbf{R}^m) = C_N(X, \mathbf{R}^m) \text{ or } C(X, \mathbf{C}^m) = C_N(X, \mathbf{C}^m)$$

of continuous mappings from X into \mathbf{R}^m or \mathbf{C}^m , as appropriate, with respect to $d_N(\cdot, \cdot)$ on the range. This is the same as the space of continuous mappings from X into \mathbf{R}^m or \mathbf{C}^m , as appropriate, with respect to the standard Euclidean metric on the range, because of the usual comparisons between N and the standard Euclidean norm. Equivalently, one can check that a mapping f from X into \mathbf{R}^m or \mathbf{C}^m is continuous if and only if f_j is continuous as a real or complex-valued function on X , as appropriate, for each $j = 1, \dots, m$. In particular, the space of continuous mappings from X into \mathbf{R}^m or \mathbf{C}^m is a linear subspace of $c(X, \mathbf{R}^m)$ or $c(X, \mathbf{C}^m)$, as appropriate.

Similarly, we get the space

$$(A.8.10) \quad C_b(X, \mathbf{R}^m) = C_{b, N}(X, \mathbf{R}^m) \text{ or } C_b(X, \mathbf{C}^m) = C_{b, N}(X, \mathbf{C}^m)$$

of bounded continuous mappings from X into \mathbf{R}^m or \mathbf{C}^m , as appropriate, with respect to $d_N(\cdot, \cdot)$ on the range. This is the same as the analogous space defined using the standard Euclidean metric on the range, as before. This is a linear subspace of each of (A.8.3) and (A.8.9), as appropriate.

A.9 Some remarks about n -dimensional volume

Let n be a positive integer. If E is a reasonably nice subset of \mathbf{R}^n , then the n -dimensional volume

$$(A.9.1) \quad \text{Vol}_n(E)$$

of E may be defined in the usual way. In fact, if n -dimensional volume on \mathbf{R}^n is interpreted as n -dimensional Lebesgue measure, then this is defined for Lebesgue measurable subsets of \mathbf{R}^n . This includes all open and closed subsets of \mathbf{R}^n , as well as their countable unions and intersections. Note that $\text{Vol}_n(E)$ may be $+\infty$ when E is not bounded.

If n -dimensional volume on \mathbf{R}^n is interpreted as Lebesgue outer measure, then $\text{Vol}_n(E)$ is defined for all subsets E of \mathbf{R}^n . However, this may not always behave as one might expect in terms of additivity of volumes of unions of disjoint subsets of \mathbf{R}^n . This is related to the Banach–Tarski paradox, as in [234, 239, 245].

A.9.1 Volumes of cells

Let a_1, \dots, a_n and b_1, \dots, b_n be real numbers with $a_j \leq b_j$ for each $j = 1, \dots, n$. The set

$$(A.9.2) \quad \mathcal{C} = \{x \in \mathbf{R}^n : a_j \leq x_j \leq b_j \text{ for each } j = 1, \dots, n\}$$

may be called a *cell* in \mathbf{R}^n , as on p31 of [210]. This is the same as the Cartesian product of the closed intervals $[a_j, b_j]$, $j = 1, \dots, n$. In this case,

$$(A.9.3) \quad \text{Vol}_n(\mathcal{C}) = \prod_{j=1}^n (b_j - a_j).$$

In particular, this is equal to 0 when $a_j = b_j$ for any j .

A.9.2 Translations and dilations

If $a \in \mathbf{R}^n$ and $E \subseteq \mathbf{R}^n$, then put

$$(A.9.4) \quad E + a = \{x + a : x \in E\},$$

which is the translation of E by a in \mathbf{R}^n . It is well known that

$$(A.9.5) \quad \text{Vol}_n(E + a) = \text{Vol}_n(E),$$

which is to say that the n -dimensional volume on \mathbf{R}^n is invariant under translations. It is easy to check directly that this is compatible with (A.9.3).

If $t \in \mathbf{R}$ and $E \subseteq \mathbf{R}^n$, then put

$$(A.9.6) \quad tE = \{tx : x \in E\},$$

which corresponds to dilating E by t in \mathbf{R}^n . It is also well known that

$$(A.9.7) \quad \text{Vol}_n(tE) = |t|^n \text{Vol}_n(E),$$

where the right side is interpreted as being 0 when $t = 0$, even if $\text{Vol}_n(E) = +\infty$, as mentioned in Subsection 7.6.1. One can check directly that this is compatible with (A.9.3) as well. In fact, if A is any linear mapping from \mathbf{R}^n into itself, then it is well known that $\text{Vol}_n(A(E))$ is equal to $|\det A|$ times $\text{Vol}_n(E)$, as mentioned in Subsection 7.6.1. This product is interpreted as being equal to 0 when $\det A = 0$, even if $\text{Vol}_n(E) = +\infty$, as before.

A.10 Volumes and Lipschitz mappings

Let n be a positive integer, and let N be a norm on \mathbf{R}^n , so that $d_N(x, y) = N(x - y)$ defines a metric on \mathbf{R}^n , as usual. Observe that

$$(A.10.1) \quad d_N(x + a, y + a) = N((x + a) - (y + a)) = N(x - y) = d_N(x, y)$$

for all $a, x, y \in \mathbf{R}^n$, so that $d_N(\cdot, \cdot)$ is invariant under translations on \mathbf{R}^n .

Similarly,

$$(A.10.2) \quad \begin{aligned} d_N(tx, ty) &= N(tx - ty) = N(t(x - y)) \\ &= |t| N(x - y) = |t| d_N(x, y) \end{aligned}$$

for every $x, y \in \mathbf{R}^n$ and $t \in \mathbf{R}$.

A.10.1 A basic estimate for volumes

Suppose that f is a Lipschitz mapping from $E \subseteq \mathbf{R}^n$ into \mathbf{R}^n with constant $C \geq 0$ with respect to $d_N(\cdot, \cdot)$ and its restriction to E , so that

$$(A.10.3) \quad d_N(f(x), f(y)) \leq C d_N(x, y)$$

for every $x, y \in E$. It is well known that

$$(A.10.4) \quad \text{Vol}_n(f(E)) \leq C^n \text{Vol}_n(E).$$

One can look at this in terms of the standard way of changing variables in n -dimensional integrals, under suitable conditions. Although the latter is discussed in many textbooks, one may also be interested in the articles [35, 155, 156, 230], as well as [73, 247, 248], in connection with Lebesgue measure and integration.

Of course, Lebesgue measure and integration are discussed in many textbooks too, and one may be interested in the articles [59, 86, 157] as well.

Alternatively, one can consider (A.10.4) in terms of n -dimensional Hausdorff measure on \mathbf{R}^n with respect to $d_N(\cdot, \cdot)$. It is well known that this is equal to a constant multiple of Lebesgue outer measure on \mathbf{R}^n .

Note that a version of (A.10.4) for linear mappings on \mathbf{R}^n is discussed in Subsection 7.6.2.

If α is any positive real number, then α -dimensional Hausdorff measure may be defined on any metric space. There is an analogue of (A.10.4) for Lipschitz

mappings between arbitrary metric spaces, using Hausdorff measures of the same dimension on the domain and range.

If k is a positive integer less than or equal to n , then k -dimensional Hausdorff measure on \mathbf{R}^n with respect to the standard Euclidean metric is related to the usual k -dimensional volume of reasonably nice k -dimensional submanifolds of \mathbf{R}^n .

A.11 Compact support

Let $(X, d(x, y))$ be a metric space, and let f be a real or complex-valued function on X . The *support* of f in X is defined by

$$(A.11.1) \quad \text{supp } f = \overline{\{x \in X : f(x) \neq 0\}},$$

which is to say the closure of the set of $x \in X$ with $f(x) \neq 0$ with respect to $d(\cdot, \cdot)$. Note that this is different from the definition of the support used in Section 1.6.

Suppose for the moment that $d(x, y)$ is the *discrete metric* on X , which is equal to 1 when $x \neq y$, and to 0 when $x = y$. In this case, it is easy to see that

$$(A.11.2) \quad \text{every subset of } X \text{ is a closed set.}$$

Similarly, the closure of any subset E of X is the same as E . This means that (A.11.1) is the same as the previous definition of the support in Section 1.6 under these conditions.

A.11.1 Some properties of the support

If $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, then

$$(A.11.3) \quad \begin{aligned} \{x \in X : t f(x) \neq 0\} &= \{x \in X : f(x) \neq 0\} && \text{when } t \neq 0 \\ &= \emptyset && \text{when } t = 0. \end{aligned}$$

This implies that

$$(A.11.4) \quad \begin{aligned} \text{supp}(t f) &= \text{supp } f && \text{when } t \neq 0 \\ &= \emptyset && \text{when } t = 0. \end{aligned}$$

If g is another real or complex-valued function on X , then it is easy to see that

$$(A.11.5) \quad \{x \in X : f(x) + g(x) \neq 0\} \subseteq \{x \in X : f(x) \neq 0\} \cup \{x \in X : g(x) \neq 0\}.$$

This implies that

$$(A.11.6) \quad \text{supp}(f + g) \subseteq (\text{supp } f) \cup (\text{supp } g).$$

More precisely, this also uses the well-known fact that

$$(A.11.7) \quad \overline{(E_1 \cup E_2)} \subseteq \overline{E_1} \cup \overline{E_2}$$

for any two subsets E_1, E_2 of X .

A.11.2 Compactly supported functions

Of course, f is said to have *compact support* in X if the support of f is a compact subset of X . Suppose that there is a compact subset K of X such that

$$(A.11.8) \quad \{x \in X : f(x) \neq 0\} \subseteq K.$$

This implies that

$$(A.11.9) \quad \text{supp } f \subseteq K,$$

because compact subsets of metric spaces are closed sets. If E is a closed subset of X such that $E \subseteq K$, then it is well known that E is also compact in X . Thus (A.11.9) implies that $\text{supp } f$ is compact, because $\text{supp } f$ is a closed set in X , by construction.

If f has compact support in X , then tf has compact support in X for every $t \in \mathbf{R}$ or \mathbf{C} , as appropriate, because of (A.11.4).

Let g be another real or complex-valued function on X , as appropriate, with compact support. It is well known and easy to see that the union of any two compact subsets of X is also compact. This means that the union of $\text{supp } f$ and $\text{supp } g$ is compact. Thus (A.11.6) implies that the support of $f + g$ is contained in a compact subset of X . It follows that the support of $f + g$ is compact, as in the preceding paragraph.

A.11.3 Compactly supported continuous functions

Suppose that $X \neq \emptyset$, and let $C_{com}(X, \mathbf{R})$ and $C_{com}(X, \mathbf{C})$ be the spaces of continuous real and complex-valued functions on X with compact support, respectively. These are linear subspaces of the real and complex vector spaces $C(X, \mathbf{R})$ and $C(X, \mathbf{C})$ of continuous real and complex-valued functions on X , respectively, by the remarks in the previous paragraphs. Of course, if X is compact, then every real or complex-valued function on X has compact support.

Suppose that f is a continuous real or complex-valued function on X with compact support. This implies that

$$(A.11.10) \quad f(\text{supp } f)$$

is a compact subset of \mathbf{R} or \mathbf{C} , as appropriate, because continuous mappings send compact sets to compact sets. In particular, this means that f is bounded on $\text{supp } f$, because compact subsets of metric spaces are bounded. It follows that f is bounded on X , because f is equal to 0 on the complement of $\text{supp } f$, by construction. Thus

$$(A.11.11) \quad C_{com}(X, \mathbf{R}) \subseteq C_b(X, \mathbf{R}), \quad C_{com}(X, \mathbf{C}) \subseteq C_b(X, \mathbf{C}),$$

where $C_b(X, Y)$ is as in Section 1.14.

Suppose for the moment again that X is equipped with the discrete metric. In this case, the only compact subsets of X are those with only finitely many elements. If Y is any metric space, then every mapping from X into Y is continuous. Hence $C_{com}(X, \mathbf{R})$ and $C_{com}(X, \mathbf{C})$ are the same as the spaces $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$ defined in Section 1.6 in this situation.

A.12 Local compactness

A metric space (X, d) is said to be *locally compact* at a point $x \in X$ if there is a positive real number r such that

$$(A.12.1) \quad \overline{B}(x, r) \text{ is a compact subset of } X.$$

Remember that $\overline{B}(x, r)$ is the closed ball in X centered at x with radius r , and that this is a closed set in X , as mentioned in Subsection 1.9.3. If X is locally compact at x , then (A.12.1) holds for all sufficiently small $r > 0$. This uses the fact that if a closed set in X is contained in a compact subset of X , then that closed set is compact as well.

If X is locally compact at every $x \in X$, then X is said to be *locally compact* as a metric space. Of course, if X is compact, then X is automatically locally compact. If X is equipped with the discrete metric, then X is clearly locally compact. Note that \mathbf{R}^n is locally compact with respect to the standard Euclidean metric for every positive integer n .

Of course, if E is any subset of X , then the restriction of $d(\cdot, \cdot)$ defines a metric on E . If $x \in E$ and E is a closed set in X , then one can check that E is locally compact at x too. The same conclusion holds when E is an open set in X . This uses the well-known fact that a subset of E is compact as a subset of E if and only if it is compact as a subset of X .

A.12.1 Local compactness and continuous functions

If there is a continuous real or complex-valued function f on X with compact support such that $f(x) \neq 0$, then X is locally compact at x . More precisely, if $f(x) \neq 0$, then there is an $r > 0$ such that f is nonzero at every point in $\overline{B}(x, r)$, because f is continuous at x . This implies that

$$(A.12.2) \quad \overline{B}(x, r) \subseteq \text{supp } f.$$

It follows that $\overline{B}(x, r)$ is compact when $\text{supp } f$ is compact, because $\overline{B}(x, r)$ is a closed set in X .

If (X, d) is any metric space and $x \in X$, then it is well known and not difficult to show that the distance to x defines a continuous real-valued function on X . If r is a positive real number, then one can use this to get continuous nonnegative real-valued functions f on X such that $f(x) > 0$ and

$$(A.12.3) \quad \text{supp } f \subseteq \overline{B}(x, r).$$

More precisely, one can obtain such functions f by composing the distance to x with suitable continuous functions on the real line. If X is locally compact at x , then one can choose $r > 0$ small enough so that $\overline{B}(x, r)$ is compact in X . In this case, (A.12.3) implies that f has compact support in X , as in Subsection A.11.2.

A.13 Another vanishing condition

Let $(X, d(x, y))$ be a metric space, and let f be a real or complex-valued function on X . We say that f *vanishes at infinity* on X if for every $\epsilon > 0$ there is a compact subset $K(\epsilon)$ of X such that

$$(A.13.1) \quad |f(x)| < \epsilon$$

for every $x \in X \setminus K(\epsilon)$.

Note that this is not the same as the definition used in Section 2.5. However, if X is equipped with the discrete metric, then the two notions are equivalent, because the only compact subsets of X are those with only finitely many elements. If X is a compact metric space, then any real or complex-valued function on X automatically vanishes at infinity in this sense.

A.13.1 Some equivalent reformulations

Equivalently, f vanishes at infinity on X if for every $\epsilon > 0$ there is a compact set $K(\epsilon) \subseteq X$ such that

$$(A.13.2) \quad E_\epsilon(f) = \{x \in X : |f(x)| \geq \epsilon\} \subseteq K(\epsilon).$$

This condition implies that

$$(A.13.3) \quad \overline{E_\epsilon(f)} \subseteq K(\epsilon),$$

where $\overline{E_\epsilon(f)}$ is the closure of $E_\epsilon(f)$ in X , because compact subsets of metric spaces are closed sets. It follows from (A.13.3) that

$$(A.13.4) \quad \overline{E_\epsilon(f)} \text{ is compact in } X,$$

because a closed set in a metric space is compact when it is contained in a compact set.

Conversely, if (A.13.4) holds, then one might as well take

$$(A.13.5) \quad K(\epsilon) = \overline{E_\epsilon(f)}.$$

Thus f vanishes at infinity on X if and only if (A.13.4) holds for every $\epsilon > 0$.

If f is continuous on X , then one can check that

$$(A.13.6) \quad E_\epsilon(f) \text{ is a closed set in } X$$

for every $\epsilon > 0$. This means that $\overline{E_\epsilon(f)} = E_\epsilon(f)$ for every $\epsilon > 0$, so that f vanishes at infinity on X if and only if

$$(A.13.7) \quad E_\epsilon(f) \text{ is compact in } X$$

for every $\epsilon > 0$.

Note that

$$(A.13.8) \quad \overline{E_\epsilon(f)} \subseteq \text{supp } f$$

for every $\epsilon > 0$, where $\text{supp } f$ is as in (A.11.1). If f has compact support in X , then it follows that f vanishes at infinity on X .

A.13.2 Continuous functions vanishing at infinity

If f vanishes at infinity on X , then it is easy to see that tf vanishes at infinity for every $t \in \mathbf{R}$ or \mathbf{C} , as appropriate.

Let g be another real or complex-valued function on X , as appropriate. If f and g both vanish at infinity on X , then one can check that $f + g$ vanishes at infinity on X too. This uses the fact that the union of two compact subsets of X is compact as well, as mentioned in Subsection A.11.2.

Suppose that $X \neq \emptyset$, and let $C_0(X, \mathbf{R})$ and $C_0(X, \mathbf{C})$ be the spaces of continuous real and complex-valued functions on X that vanish at infinity, respectively. These are linear subspaces of the real and complex vector spaces $C(X, \mathbf{R})$ and $C(X, \mathbf{C})$ of continuous real and complex-valued functions on X , respectively, by the remarks in the preceding paragraph. We also have that

$$(A.13.9) \quad C_{com}(X, \mathbf{R}) \subseteq C_0(X, \mathbf{R}), \quad C_{com}(X, \mathbf{C}) \subseteq C_0(X, \mathbf{C})$$

because functions on X with compact support automatically vanish at infinity, as before. If X is equipped with the discrete metric, then $C_0(X, \mathbf{R})$ and $C_0(X, \mathbf{C})$ are the same as the spaces $c_{00}(X, \mathbf{R})$ and $c_{00}(X, \mathbf{C})$ defined in Section 2.5, respectively. This uses the fact that all functions on X are continuous in this case.

Let f be a continuous real or complex-valued function on X . Of course, the restriction of f to any compact set $K \subseteq X$ is bounded, because $f(K)$ is compact, and hence bounded. If (A.13.7) holds for any $\epsilon > 0$, then it follows that f is bounded on X . In particular, if f vanishes at infinity on X , then f is bounded on X . This means that

$$(A.13.10) \quad C_0(X, \mathbf{R}) \subseteq C_b(X, \mathbf{R}), \quad C_0(X, \mathbf{C}) \subseteq C_b(X, \mathbf{C}),$$

where $C_b(X, Y)$ is as in Section 1.14.

Let f be a continuous real or complex-valued function on X again, and let x be an element of X such that $f(x) \neq 0$. Thus $|f(x)| > \epsilon$ for some $\epsilon > 0$, and the continuity of f at x implies that there is an $r > 0$ such that $|f(y)| \geq \epsilon$ for every $y \in X$ with $d(x, y) \leq r$. Equivalently, this means that

$$(A.13.11) \quad \overline{B}(x, r) \subseteq E_\epsilon(f),$$

where $\overline{B}(x, r)$ is the closed ball in X centered at x with radius r , as before. If f vanishes at infinity on X , then it follows that $\overline{B}(x, r)$ is compact, because $\overline{B}(x, r)$ is a closed set in X . This implies that X is locally compact at x under these conditions.

A.13.3 A uniform convergence property

Let $\{f_j\}_{j=1}^\infty$ be a sequence of real or complex-valued functions on X that converges uniformly to a real or complex-valued function f on X , as appropriate. Suppose that f_j vanishes at infinity on X for each j , and let us show that f vanishes at infinity on X too.

Let $\epsilon > 0$ be given. Because $\{f_j\}_{j=1}^\infty$ converges to f uniformly on X , there is an $L \in \mathbf{Z}_+$ such that

$$(A.13.12) \quad |f_j(x) - f(x)| < \epsilon/2$$

for every $x \in X$ and $j \geq L$. There is also a compact set $K \subseteq X$ such that

$$(A.13.13) \quad |f_L(x)| < \epsilon/2$$

for every $x \in X \setminus K$, because f_L vanishes at infinity on X . Combining (A.13.12) and (A.13.13), we get that

$$(A.13.14) \quad |f(x)| \leq |f(x) - f_L(x)| + |f_L(x)| < \epsilon/2 + \epsilon/2 = \epsilon$$

for every $x \in X \setminus K$, as desired.

In particular, this implies that $C_0(X, \mathbf{R})$ and $C_0(X, \mathbf{C})$ are closed sets in $C_b(X, \mathbf{R})$ and $C_b(X, \mathbf{C})$ with respect to the supremum metric.

A.13.4 Some additional properties

Let f be a real or complex-valued function on X that vanishes at infinity. Also let ϕ be a mapping from \mathbf{R} or \mathbf{C} into itself such that $\phi(0) = 0$ and ϕ is continuous at 0. Under these conditions, one can check that

$$(A.13.15) \quad \phi \circ f \text{ vanishes at infinity on } X$$

too.

If ϕ vanishes on a neighborhood of 0 in \mathbf{R} or \mathbf{C} , as appropriate, then one can verify that

$$(A.13.16) \quad \phi \circ f \text{ has compact support in } X.$$

Of course, if f and ϕ are continuous, then $\phi \circ f$ is continuous on X too. It is easy to see that the identity mappings on \mathbf{R} and \mathbf{C} can be approximated uniformly by continuous functions that vanish on a neighborhood of 0. If f is a continuous function on X that vanishes at infinity, then one can use this to approximate f uniformly by continuous functions on X with compact support.

This means that $C_{com}(X, \mathbf{R})$ and $C_{com}(X, \mathbf{C})$ are dense in $C_0(X, \mathbf{R})$ and $C_0(X, \mathbf{C})$, respectively, with respect to the supremum metric.

Bibliography

- [1] S. Abian and A. Brown, *On the solution of the differential equation $f(x, y, y') = 0$* , American Mathematical Monthly **66** (1959), 192–199.
- [2] R. Agarwal, C. Flaut, and D. O'Regan, *An Introduction to Real Analysis*, CRC Press, 2018.
- [3] T. Apostol, *Mathematical Analysis*, 2nd edition, Addison-Wesley, 1974.
- [4] T. Archibald, *Connectivity and smoke-rings: Green's second identity in its first fifty years*, Mathematics Magazine **62** (1989), 219–232.
- [5] J. Ash, *Uniqueness of representation by trigonometric series*, American Mathematical Monthly **96** (1989), 873–885.
- [6] G. Aumann, *Classroom notes: The Lipschitz condition and uniqueness*, American Mathematical Monthly **72** (1965), 652–653.
- [7] L. Badger, *Notes: Generating the measures of n -balls*, American Mathematical Monthly **107** (2000), 256–258.
- [8] R. Bagley and J. McKnight, *Differentiation of infinite series of functions*, American Mathematical Monthly **66** (1959), 890–892.
- [9] J. Baker, *Isometries in normed spaces*, American Mathematical Monthly **78** (1971), 655–658.
- [10] J. Baker, *Integration over spheres and the divergence theorem for balls*, American Mathematical Monthly **104** (1997), 36–47.
- [11] J. Baker, *Integration of radial functions*, Mathematics Magazine **72** (1999), 392–395.
- [12] P. Barry and D. Hurley, *On series of partial fractions*, American Mathematical Monthly **98** (1991), 240–242.
- [13] R. Beals, *Advanced Mathematical Analysis: Periodic Functions and Distributions, Complex Analysis, Laplace Transform and Applications*, Springer-Verlag, 1973.

- [14] R. Beals, *Analysis: An Introduction*, Cambridge University Press, 2004.
- [15] R. Beals, *The scope of the power series method*, *Mathematics Magazine* **86** (2013), 56–62.
- [16] P. Beesack, *On improper multiple integrals*, *Mathematics Magazine* **43** (1970), 113–123.
- [17] M. Berger, *Convexity*, *American Mathematical Monthly* **97** (1990), 650–678.
- [18] L. Bers, *Classroom notes: On avoiding the mean value theorem*, *American Mathematical Monthly* **74** (1967), 583.
- [19] Á. Besenyei, *Lebesgue's road to antiderivatives*, *Mathematics Magazine* **86** (2013), 255–260.
- [20] Á. Besenyei, *Peano's unnoticed proof of Borel's theorem*, *American Mathematical Monthly* **121** (2014), 69–72.
- [21] R. Boas, Jr., *Partial sums of infinite series, and how they grow*, *American Mathematical Monthly* **84** (1977), 237–258.
- [22] R. Boas, Jr., *Lion Hunting & Other Mathematical Pursuits: A Collection of Mathematics, Verse, and Stories*, Edited and with an introduction by G. Alexanderson and D. Mugler, Mathematical Association of America, 1995.
- [23] R. Boas, Jr., *A Primer of Real Functions*, 4th edition, revised and updated by H. Boas, Mathematical Association of America, 1996.
- [24] S. Bochner, *Fourier series came first*, *American Mathematical Monthly* **86** (1979), 197–199.
- [25] J. Borzellino and M. Sherman, *When is a trigonometric polynomial not a trigonometric polynomial?*, *American Mathematical Monthly* **119** (2012), 422–425.
- [26] M. Botsko and R. Gosser, *The teaching of mathematics: On the differentiability of functions of several variables*, *American Mathematical Monthly* **92** (1985), 663–665.
- [27] D. Bressoud, *A Radical Approach to Real Analysis*, 2nd edition, Mathematical Association of America, 2007.
- [28] D. Bressoud, *A Radical Approach to Lebesgue's Theory of Integration*, MAA Textbooks, Cambridge University Press, 2008.
- [29] D. Bressoud, *Historical reflections on teaching the fundamental theorem of integral calculus*, *American Mathematical Monthly* **118** (2011), 99–115.

- [30] D. Bressoud, *Calculus Reordered: A History of the Big Ideas*, Princeton University Press, 2019.
- [31] A. Browder, *Mathematical Analysis: An Introduction*, Springer-Verlag, 1996.
- [32] A. Browder, *Topology in the complex plane*, American Mathematical Monthly **107** (2000), 393–401.
- [33] A. Browder, *Complex numbers and the ham sandwich theorem*, American Mathematical Monthly **113** (2006), 935–937.
- [34] A. Brown, *A proof of the Lebesgue condition for Riemann integrability*, American Mathematical Monthly **43** (1936), 396–398.
- [35] A. Brown, *On transformation of multiple integrals*, American Mathematical Monthly **48** (1941), 29–33.
- [36] A. Brown, *Mathematical notes: Extensions of the Brouwer fixed point theorem*, American Mathematical Monthly **69** (1962), 643.
- [37] A. Bruckner, *Derivatives: why they elude classification*, Mathematics Magazine **49** (1976), 5–11.
- [38] G. Bullock, *The teaching of mathematics: A geometric interpretation of the Riemann–Stieltjes integral*, American Mathematical Monthly **95** (1988), 448–455.
- [39] R. Burckel and C. Goffman, *Rectifiable curves are of zero content*, Mathematics Magazine **44** (1971), 179–180.
- [40] J. Burkill and H. Burkill, *A Second Course in Mathematical Analysis*, Cambridge University Press, 1970.
- [41] T. Cameron, *The determinant from signed volume to the Laplace expansion*, American Mathematical Monthly **126** (2019), 437–447.
- [42] D. Cannell, *Geoge Green: an enigmatic mathematician*, American Mathematical Monthly **106** (1999), 136–151.
- [43] D. Cannell, *George Green, Mathematician & Physicist 1793–1841*, 2nd edition, Society for Industrial and Applied Mathematics (SIAM), 2001.
- [44] P. Chernoff, *Pointwise convergence of Fourier series*, American Mathematical Monthly **87** (1980), 399–400; letter to the editor, **92** (1985), 373–374.
- [45] J. Cid and R. López Pouso, *Does Lipschitz with respect to x imply uniqueness for the differential equation $y' = f(x, y)$?*, American Mathematical Monthly **116** (2009), 61–66.

- [46] J. Cid and R. López Pouso, *Integration by parts and by substitution unified with applications to Green's theorem and uniqueness for ODEs*, American Mathematical Monthly **123** (2016), 40–52.
- [47] L. Cohen, *Classroom notes: On being mean to the mean-value theorem*, American Mathematical Monthly **74** (1967), 581–582.
- [48] R. Cohen, *Set isometries and their extensions in the max metric*, American Mathematical Monthly **97** (1990), 795–808.
- [49] G. Convertito and D. Cruz-Urbe, OFS, *The Stieltjes integral*, CRC Press, 2023.
- [50] J. Conway, *A First Course in Analysis*, Cambridge University Press, 2018.
- [51] R. Cooke, *The Cantor–Lebesgue theorem*, American Mathematical Monthly **86** (1979), 558–565.
- [52] D. Cruz-Urbe, SFO and C. Neugebauer, *An elementary proof of the error estimates for the trapezoidal rule*, Mathematics Magazine **76** (2003), 303–306.
- [53] X. Dai, *Continuous differentiability of solutions of ODEs with respect to initial conditions*, American Mathematical Monthly **113** (2006), 66–70.
- [54] N. de Silva, *A concise, elementary proof of Arzelà's bounded convergence theorem*, American Mathematical Monthly **117** (2010), 918–920.
- [55] J. Dettman, *The solution of a second order linear differential equation near a regular singular point*, American Mathematical Monthly **71** (1964), 378–385.
- [56] J. Dettman, *Power series solutions of ordinary differential equations*, American Mathematical Monthly **74** (1967), 428–430.
- [57] A. Devinatz, *Advanced Calculus*, Holt, Rinehart and Winston, 1968.
- [58] T. Drammatico, *A new proof of the equality of mixed second partial derivatives*, American Mathematical Monthly **122** (2015), 602–603.
- [59] R. Dressler and K. Stromberg, *The Tonelli integral*, American Mathematical Monthly **81** (1974), 67–68.
- [60] L. Dubins, *On differentiation of series term-by-term*, American Mathematical Monthly **67** (1960), 771–772.
- [61] J. Duemmel, *Classroom notes: Continuous dependence on parameters*, American Mathematical Monthly **76** (1969), 294–295.
- [62] W. Dunham, *Touring the calculus gallery*, American Mathematical Monthly **112** (2005), 1–19.

- [63] W. Dunham, *The Calculus Gallery: Masterpieces from Newton to Lebesgue*, Princeton University Press, 2018.
- [64] P. Duren, *Invitation to Classical Analysis*, American Mathematical Society, 2012.
- [65] J. Fabrey, *Picard's theorem*, American Mathematical Monthly **79** (1972), 1020–1023.
- [66] A. Fadell, *Classroom notes: A proof of the chain rule for derivatives in n -space*, American Mathematical Monthly **80**, 1134–1135.
- [67] W. Felscher, *Bolzano, Cauchy, epsilon, delta*, American Mathematical Monthly **107** (2000), 844–862.
- [68] J. Fisher, *Curves of constant width from a linear viewpoint*, Mathematics Magazine **60** (1987), 131–140.
- [69] J. Fisher and J. Shilleto, *Three aspects of Fubini's theorem*, Mathematics Magazine **59** (1986), 40–42.
- [70] L. Flatto, *Advanced Calculus*, Williams & Wilkins, 1976.
- [71] W. Fleming, *Functions of Several Variables*, 2nd edition, Springer-Verlag, 1977.
- [72] H. Fletcher, *Summing of trigonometric series with coefficients which have a periodic factor*, American Mathematical Monthly **65** (1958), 349–351.
- [73] T. Flett, *On transformations in \mathbf{R}^n and a theorem of Sard*, American Mathematical Monthly **71** (1964), 623–629.
- [74] G. Folland, *Fourier Analysis and its Applications*, Wadsworth & Brooks / Cole, 1992.
- [75] G. Folland, *How to integrate a polynomial over a sphere*, American Mathematical Monthly **108** (2001), 446–448.
- [76] G. Folland, *A Guide to Advanced Real Analysis*, Mathematical Association of America, 2009.
- [77] M. Fort, Jr., *Some properties of continuous functions*, American Mathematical Monthly **59** (1952), 372–375.
- [78] J. Foster and F. Richards, *The Gibbs phenomenon for piecewise-linear approximation*, American Mathematical Monthly **98** (1991), 47–49.
- [79] P. Franklin, *A Treatise on Advanced Calculus*, Dover, 1964.
- [80] A. Friedman, *Advanced Calculus*, Holt, Rinehart and Winston, 1971.

- [81] T. Gamelin and R. Greene, *Introduction to Topology*, 2nd edition, Dover, 1999.
- [82] C. Gardner, *Another elementary proof of Peano's existence theorem*, American Mathematical Monthly **83** (1976), 556-560.
- [83] H. Gaskill and P. Narayanaswami, *Foundations of Analysis: The Theory of Limits*, Harper & Row, 1989.
- [84] B. Gelbaum and J. Olmsted, *Counterexamples in Analysis*, corrected reprint of the 2nd (1965) edition, Dover, 2003.
- [85] A. Gluchoff, *Trigonometric series and theories of integration*, Mathematics Magazine **67** (1994), 3-20.
- [86] C. Goffman, *Definition of the Lebesgue integral*, American Mathematical Monthly **60** (1953), 251-252.
- [87] C. Goffman, *Real Functions*, Rinehart & Co., 1953.
- [88] C. Goffman, *Classroom notes: Arc length*, American Mathematical Monthly **71** (1964), 303-304.
- [89] C. Goffman, *Calculus of Several Variables*, Harper & Row, 1965.
- [90] C. Goffman, *Introduction to Real Analysis*, Harper & Row, 1966.
- [91] C. Goffman, *A note on integration*, Mathematics Magazine **44** (1971), 1-4.
- [92] C. Goffman, *A bounded derivative which is not Riemann integrable*, American Mathematical Monthly **84** (1977), 205-206.
- [93] R. Goldberg, *Methods of Real Analysis*, 2nd edition, Wiley, 1976.
- [94] E. González-Velasco, *Connections in mathematical analysis: the case of Fourier series*, American Mathematical Monthly **99** (1992), 427-441.
- [95] E. González-Velasco, *James Gregory's calculus in the Geometriæ pars universalis*, American Mathematical Monthly **114** (2007), 565-576.
- [96] E. González-Velasco, *Journey Through Mathematics: Creative Episodes in its History*, Springer, 2011.
- [97] R. Gordon, *A convergence theorem for the Riemann integral*, Mathematics Magazine **73** (2000), 141-147.
- [98] R. Gordon, *A bounded derivative that is not Riemann integrable*, Mathematics Magazine **89** (2016), 364-370.
- [99] F. Gouvêa, *Was Cantor surprised?*, American Mathematical Monthly **118** (2011), 198-209.

- [100] F. Gouvêa, *A Short Book on Long Sums — Infinite Series for Calculus Students*, Springer, 2023.
- [101] J. Grabiner, *The Origins of Cauchy's Rigorous Calculus*, MIT Press, 1981.
- [102] J. Grabiner, *Who gave you the epsilon? Cauchy and the origins of rigorous calculus*, *American Mathematical Monthly* **90** (1983), 185–194.
- [103] J. Grabiner, *The changing concept of change: the derivative from Fermat to Weierstrass*, *Mathematics Magazine* **56** (1983), 195–206.
- [104] J. Grabiner, *Was Newton's calculus a dead end? The continental influence of Maclaurin's treatise of fluxions*, *American Mathematical Monthly* **104** (1997), 393–410.
- [105] J. Grabiner, *A Historian Looks Back: The Calculus as Algebra and selected writings*, MAA Spectrum, Mathematical Association of America, 2010.
- [106] I. Grattan-Guinness, *Why did George Green write his essay of 1828 on electricity and magnetism?*, *American Mathematical Monthly* **102** (1995), 387–396.
- [107] J. Gray, *Miscellanea: Clear theory*, *American Mathematical Monthly* **92** (1985), 158–159.
- [108] J. Gray, *The Real and the Complex: A History of Analysis in the 19th Century*, Springer, 2015.
- [109] J. Gray, *Change and Variations — A History of Differential Equations to 1900*, Springer, 2021.
- [110] J. Green, *The theory of functions of a real variable*, *Mathematics Magazine* **24** (1951), 209–217.
- [111] J. Green and F. Valentine, *On the Arzelà–Ascoli theorem*, *Mathematics Magazine* **34** (1960/61), 199–202.
- [112] R. Gunning, *An Introduction to Analysis*, Princeton University Press, 2018.
- [113] L. Gurin, *A problem*, *American Mathematical Monthly* **103** (1996), 683–686.
- [114] D. Hallenbeck and K. Tkaczyńska, *The absolute and uniform convergence of infinite improper integrals*, *American Mathematical Monthly* **95** (1988), 124–126.
- [115] J. Hannah, *A geometric approach to determinants*, *American Mathematical Monthly* **103** (1996), 401–409.

- [116] B. Hasselblatt, *Laws of exponentis by differentiation*, American Mathematical Monthly **129** (2022), 984.
- [117] I. Herron, *The radius of convergence of power series solutions to linear differential equations*, American Mathematical Monthly **96** (1989), 824–827.
- [118] E. Hewitt and K. Stromberg, *Real and Abstract Analysis*, Springer-Verlag, 1975.
- [119] T. Hildebrandt, *Definitions of Stieltjes integrals of the Riemann type*, American Mathematical Monthly **45** (1938), 265–278.
- [120] E. Hille, *Mathematics and mathematicians from Abel to Zermelo*, Mathematics Magazine **26** (1953), 127–146.
- [121] K. Hoffman, *Analysis in Euclidean Space*, Prentice-Hall, 1975.
- [122] I. Hirschman, Jr., *Infinite Series*, Holt, Rinehart and Winston, 1962.
- [123] P. Johnson and R. Redheffer, *Scrambled series*, American Mathematical Monthly **73** (1966), 822–828.
- [124] F. Jones, *Lebesgue Integration on Euclidean Space*, Jones and Bartlett, 1993.
- [125] L. Jones, *An elementary derivation of the numerical integration bounds in beginning calculus*, American Mathematical Monthly **124** (2017), 558–561.
- [126] H. Junghenn, *A Course in Real Analysis*, CRC Press, 2015.
- [127] W. Kaczor and M. Nowak, *Problems in Mathematical Analysis I: Real Numbers, Sequences and Series*, translated and revised from the 1996 Polish original by the authors, American Mathematical Society, 2000.
- [128] W. Kaczor and M. Nowak, *Problems in Mathematical Analysis II: Continuity and Differentiation*, translated from the 1998 Polish original, revised, and augmented by the authors, American Mathematical Society, 2001.
- [129] W. Kaczor and M. Nowak, *Problems in Mathematical Analysis III: Integration*, American Mathematical Society, 2003.
- [130] I. Kaplansky, *Set Theory and Metric Spaces*, 2nd edition, AMS Chelsea, 1977.
- [131] H. Kaptanoğlu, *In praise of $x^\alpha \sin(\frac{1}{x})$* , American Mathematical Monthly **108** (2001), 144–150.
- [132] D. Kearns, *Classroom notes: On solutions of linear differential equations*, American Mathematical Monthly **67** (1960), 369–370.

- [133] D. Kearns, *Classroom notes: Elementary uniqueness theorems for differential equations*, American Mathematical Monthly **68** (1961), 275–277.
- [134] H. Kennedy, *Research problems: Is there an elementary proof of Peano's existence theorem for first order differential equations?*, American Mathematical Monthly **76** (1969), 1043–1045.
- [135] H. Kennedy, *Peano: Life and Works of Giuseppe Peano*, Reidel, 1980.
- [136] G. Klambauer, *Mathematical Analysis*, Dekker, 1975.
- [137] A. Knapp, *Basic Real Analysis*, Birkhäuser, 2005.
- [138] J. Koliha, *Mean, meaner, and meanest mean value theorem*, American Mathematical Monthly **116** (2009), 356–361.
- [139] V. Komornik, *A short, elementary introduction to the exponential function*, American Mathematical Monthly **127** (2020), 841–846.
- [140] T. Körner, *Differentiable functions on the rationals*, Bulletin of the London Mathematical Society **23** (1991), 557–562.
- [141] T. Körner, *Exercises for Fourier Analysis* ([142]). Cambridge University Press, 1993.
- [142] T. Körner, *Butterfly hunting and the symmetry of mixed partial derivatives*, American Mathematical Monthly **105** (1998), 756–758.
- [143] T. Körner, *A Companion to Analysis: A Second First and First Second Course in Analysis*, American Mathematical Society, 2004.
- [144] T. Körner, *Fourier Analysis*, Cambridge University Press, 2022.
- [145] S. Krantz, *A Guide to Real Variables*, Mathematical Association of America, 2009.
- [146] S. Krantz, *Real Analysis and Foundations*, 4th edition, CRC Press, 2017.
- [147] S. Krantz, *Elementary Introduction to the Lebesgue Integral*, CRC Press, 2018.
- [148] S. Krantz and H. Parks, *The Implicit Function Theorem: History, Theory, and Applications*, Birkhäuser, 2013.
- [149] K. Kreith, *Geometric interpretation of the Implicit Function Theorem*, Mathematics Magazine **36** (1963), 64–65.
- [150] W. Kulpa, *The Poincaré–Miranda theorem*, American Mathematical Monthly **104** (1997), 545–550.
- [151] M. Laczkovich and V. Sós, *Real Analysis — Foundations and Functions of One Variable*, 5th edition, translated from the third (2005) Hungarian edition by the authors, Springer, 2015.

- [152] S. Lang, *Undergraduate Analysis*, 2nd edition, Springer-Verlag, 1997.
- [153] R. Langer, *Fourier's Series: The Genesis and Evolution of a Theory*, The first Herbert Ellsworth Slaughter Memorial Paper, American Mathematical Monthly **54** (1947), no. 7, part II.
- [154] J. LaVita, *Classroom notes: A necessary and sufficient condition for Riemann integration*, American Mathematical Monthly **71** (1964), 193–196.
- [155] P. Lax, *Change of variables in multiple integrals*, American Mathematical Monthly **106** (1999), 497–501.
- [156] P. Lax, *Change of variables in multiple integrals II*, American Mathematical Monthly **108** (2001), 115–119.
- [157] P. Lax, *Rethinking the Lebesgue integral*, American Mathematical Monthly **116** (2009), 863–881.
- [158] S. Lefschetz, *A direct proof of De Moivre's formula*, American Mathematical Monthly **23** (1916), 366–368.
- [159] T. Lehrer, *The Derivative Song*, American Mathematical Monthly **81** (1974), 490.
- [160] T. Lehrer, *There's a Delta for Every Epsilon (Calypso)*, American Mathematical Monthly **81** (1974), 612.
- [161] T. Lehrer, *The Professor's Song*, American Mathematical Monthly **81** (1974), 745.
- [162] R. Leipnik, *Note on alternating Tannery series*, American Mathematical Monthly **65** (1958), 197–198.
- [163] P. Leonetti, *A characterization of Cesàro convergence*, American Mathematical Monthly **128** (2021), 559–562.
- [164] L. Levine, *On a necessary and sufficient condition for Riemann integrability*, American Mathematical Monthly **84** (1977), 205.
- [165] J. Lewin, *The teaching of mathematics: A truly elementary approach to the bounded convergence theorem*, American Mathematical Monthly **93** (1986), 395–397.
- [166] J. Lewin, *The teaching of mathematics: Some applications of the bounded convergence theorem for an introductory course in analysis*, American Mathematical Monthly **94** (1987), 988–993.
- [167] C.-K. Li, *Norms, isometries, and isometry groups*, American Mathematical Monthly **107** (2000), 334–340.
- [168] C.-K. Li and W. So, *Isometries of l_p -norm*, American Mathematical Monthly **101** (1994), 452–453.

- [169] G. Lion, *A simple proof of the Dirichlet–Jordan convergence test*, American Mathematical Monthly **93** (1996), 281–282.
- [170] W. Luxemburg, *Arzelà’s dominated convergence theorem for the Riemann integral*, American Mathematical Monthly **78** (1971), 970–979.
- [171] S. Malik and S. Arora, *Mathematical Analysis*, 2nd edition, Wiley, 1992.
- [172] S. Mandelbrojt, *Obituary: Emile Picard, 1856–1941*, American Mathematical Monthly **49** (1942), 277–278.
- [173] S. Mandelbrojt, *The mathematical work of Jacques Hadamard*, American Mathematical Monthly **60** (1953), 599–604.
- [174] J. Mawhin, *Simple proofs of the Hadamard and Poincaré–Miranda theorems using the Brouwer fixed point theorem*, American Mathematical Monthly **126** (2019), 260–263.
- [175] P. McGrath, *Another proof of Clairaut’s theorem*, American Mathematical Monthly **121** (2014), 165–166.
- [176] J. McKnight, Jr., *Brown’s method of extending fixed point theorems*, American Mathematical Monthly **72** (1965), 152–155.
- [177] B. Mendelson, *Introduction to Topology*, 3rd edition, Dover, 1990.
- [178] R. Metzler, *On Riemann integrability*, American Mathematical Monthly **78** (1971), 1129–1131.
- [179] M. Meyerson, *Every power series is a Taylor series*, American Mathematical Monthly **88** (1981), 51–52.
- [180] M. Meyerson, *The x^x spindle*, Mathematics Magazine **69** (1996), 198–206.
- [181] M. Mezzino and M. Pinsky, *Leibnitz’s formula, Cauchy majorants, and linear differential equations*, Mathematics Magazine **71** (1998), 360–368.
- [182] J. Milnor, *Analytic proofs of the “hairy ball theorem” and the Brouwer fixed-point theorem*, American Mathematical Monthly **85** (1978), 521–524.
- [183] P. Mohazzabi and T. Fournelle, *Evaluation of ill-behaved power series*, American Mathematical Monthly **111** (2004), 308–321.
- [184] H. Montgomery, *Early Fourier Analysis*, American Mathematical Society, 2014.
- [185] H. Montgomery and U. Vorhauer, *Biased trigonometric polynomials*, American Mathematical Monthly **114** (2007), 804–809.
- [186] F. Morgan, *Real Analysis*, American Mathematical Society, 2005.

- [187] F. Morgan, *Real Analysis and Applications: Including Fourier Series and the Calculus of Variations*, American Mathematical Society, 2005.
- [188] C. Morrison and M. Stynes, *An intuitive proof of Brouwer's fixed point theorem in \mathbf{R}^2* , *Mathematics Magazine* **56** (1983), 38–41.
- [189] R. Mortini, *A short proof of the chain rule for differentiable mappings in \mathbf{R}^n* , *Mathematics Magazine* **85** (2012), 136–141.
- [190] S. Nadler, Jr., *A proof of Darboux's theorem*, *American Mathematical Monthly* **117** (2010), 174–175.
- [191] A. Nijenhuis, *Strong derivatives and inverse mappings*, *American Mathematical Monthly* **81** (1974), 969–980; addendum, **83** (1975), 22.
- [192] L. Olsen, *A new proof of Darboux's theorem*, *American Mathematical Monthly* **111** (2004), 713–715.
- [193] G. Peyser, *On the Cauchy–Lipschitz theorem*, *American Mathematical Monthly* **65** (1958), 760–762.
- [194] S. Ponnusamy, *Foundations of Mathematical Analysis*, Birkhäuser / Springer, 2012.
- [195] M. Protter, *Basic Elements of Real Analysis*, Springer-Verlag, 1998.
- [196] M. Protter and C. Morrey, Jr., *A First Course in Real Analysis*, 2nd edition, Springer-Verlag, 1991.
- [197] C. Pugh, *Real Mathematical Analysis*, 2nd edition, Springer, 2015.
- [198] H. Rademacher, *On the condition of Riemann integrability*, *American Mathematical Monthly* **61** (1954), 1–8.
- [199] R. Redheffer, *An elementary proof of sharp uniqueness theorems for second-order initial-value problems*, *American Mathematical Monthly* **81** (1974), 377–379.
- [200] M. Reed, *Fundamental Ideas in Analysis*, Wiley, 1998.
- [201] H. Reiter and A. Reiter, *The space of closed subsets of a convergent sequence*, *Mathematics Magazine* **69** (1996), 217–221.
- [202] A. Robert, *Fourier series of polygons*, *American Mathematical Monthly* **101** (1994), 420–428.
- [203] C. Rogers, *A less strange version of Milnor's proof of Brouwer's fixed-point theorem*, *American Mathematical Monthly* **87** (1980), 525–527.
- [204] M. Rosenlicht, *Introduction to Analysis*, Dover, 1986.

- [205] C. Rosentrater, *Varieties of Integration*, Mathematical Association of America, 2015.
- [206] K. Ross, *Another approach to Riemann–Stieltjes integrals*, American Mathematical Monthly **87** (1980), 660–662.
- [207] K. Ross, *Elementary Analysis: The Theory of Calculus*, 2nd edition, in collaboration with J. López, Springer, 2013.
- [208] W. Rudin, *Functional Analysis*, McGraw-Hill, 1973. A second edition was published in 1991.
- [209] W. Rudin, *Real and Complex Analysis*, 2nd edition, McGraw-Hill, 1974. A third edition was published in 1987.
- [210] W. Rudin, *Principles of Mathematical Analysis*, 3rd edition, McGraw-Hill, 1976.
- [211] W. Rudin, *Nonuniqueness and growth in first-order differential equations*, American Mathematical Monthly **89** (1982), 241–244.
- [212] D. Sanderson, *Advanced plane topology from an elementary standpoint*, Mathematics Magazine (1980), 81–89.
- [213] A. Sard, *Questions, discussions, and notes: Iterated and double limits*, American Mathematical Monthly **45** (1938), 375–376.
- [214] A. Sard, *Discussions and notes: A theorem on improper integrals*, American Mathematical Monthly **49** (1942), 536–537.
- [215] U. Satyanarayana, *A note on Riemann–Stieltjes integrals*, American Mathematical Monthly **87** (1980), 477–478.
- [216] D. Schattschneider, *A multiplicative metric*, Mathematics Magazine **49** (1976), 203–205.
- [217] D. Schattschneider, *The taxicab group*, American Mathematical Monthly **91** (1984), 423–428.
- [218] R. Seeley, *Classroom notes: Fubini implies Leibniz implies $F_{yx} = F_{xy}$* , American Mathematical Monthly **68** (1961), 56–57.
- [219] R. Shakarchi, *Problems and Solutions for Undergraduate Analysis ([152])*, Springer-Verlag, 1998.
- [220] I. Sheffer, *Convergence of multiply-infinite series*, American Mathematical Monthly **52** (1945), 365–376.
- [221] I. Sheffer, *Classroom notes: Note on multiple power series*, American Mathematical Monthly **70** (1963), 309–313.

- [222] D. Shelupsky, *Classroom notes: Derivation of the Gibbs phenomenon*, American Mathematical Monthly **87** (1980), 210–212.
- [223] O. Shisha, *An approach to Darboux-Stieltjes integration*, American Mathematical Monthly **72** (1965), 890–892.
- [224] H. Siegborg, *Some historical remarks concerning degree theory*, American Mathematical Monthly **88** (1981), 125–139.
- [225] J. Singh, *A noninductive proof of de Moivre's formula*, American Mathematical Monthly **125** (2018), 80.
- [226] D. Smith and M. Vamanamurthy, *How small is a unit ball?*, Mathematics Magazine **62** (1989), 101–107.
- [227] E. Stein and R. Shakarchi, *Fourier Analysis: An Introduction*, Princeton University Press, 2003.
- [228] E. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, Princeton University Press, 2005.
- [229] P. Stein, *Classroom notes: A note on the volume of a simplex*, American Mathematical Monthly **73** (1966), 299–301.
- [230] P. Stein, *A proof of a classical theorem in multiple integration*, American Mathematical Monthly **75** (1968), 160–163.
- [231] G. Stoica, *Components used in the Stone–Weierstrass theorem*, American Mathematical Monthly **127** (2020), 658.
- [232] A. Strauss, *Mathematical notes: Continuous dependence of solutions of ordinary differential equations*, American Mathematical Monthly **71** (1964), 649–652.
- [233] R. Strichartz, *The Way of Analysis*, Revised edition, Jones and Bartlett, 2000.
- [234] K. Stromberg, *The Banach–Tarski paradox*, American Mathematical Monthly **86** (1979), 151–161.
- [235] K. Stromberg, *An Introduction to Classical Real Analysis*, AMS Chelsea, 1981.
- [236] L. Talman, *Simpson's rule is exact for quintics*, American Mathematical Monthly **113** (2006), 144–155.
- [237] A. Taylor, *A note on an inequality for integrals*, American Mathematical Monthly **57** (1950), 93–96.
- [238] A. Taylor and W. Mann, *Advanced Calculus*, 3rd edition, Wiley, 1983.

- [239] A. Taylor and S. Wagon, *A paradox arising from the elimination of a paradox*, American Mathematical Monthly **126** (2019), 306–318.
- [240] H. ter Horst, *Riemann–Stieltjes and Lebesgue–Stieltjes integrability*, American Mathematical Monthly **91** (1984), 551–559.
- [241] A. Thompson, *Minkowski Geometry*, Cambridge University Press, 1996.
- [242] B. Thomson, *Monotone convergence for the Riemann integral*, American Mathematical Monthly **117** (2010), 547–550.
- [243] B. Thomson, *The bounded convergence theorem*, American Mathematical Monthly **127** (2020), 483–503.
- [244] T. Tokieda, *A mean value theorem*, American Mathematical Monthly **106** (1999), 673–674.
- [245] G. Tomkowicz and S. Wagon, *The Banach–Tarski Paradox*, 2nd edition, Cambridge University Press, 2016.
- [246] J. Väisälä, *A proof of the Mazur–Ulam theorem*, American Mathematical Monthly **110** (2003), 633–635.
- [247] D. Varberg, *On differentiable transformations in \mathbf{R}^n* , American Mathematical Monthly **73** (1966), no. 4, part II, 111–114.
- [248] D. Varberg, *Change of variables in multiple integrals*, American Mathematical Monthly **78** (1971), 42–45.
- [249] C. Wagner, *A generic approach to iterative methods*, Mathematics Magazine **55** (1982), 259–273.
- [250] J. Walter, *On elementary proofs of Peano’s existence theorems*, American Mathematical Monthly **80** (1973), 282–286.
- [251] W. Walter, *There is an elementary proof of Peano’s existence theorem*, American Mathematical Monthly **78** (1971), 170–173.
- [252] W. Walter, *Ordinary Differential Equations*, translated from the sixth German (1996) edition by R. Thompson, Springer-Verlag, 1998.
- [253] R. Weinstock, *On continuous differentiability*, American Mathematical Monthly **64** (1957), 492.
- [254] R. Weinstock, *Elementary evaluations of $\int_0^\infty e^{-x^2} dx$, $\int_0^\infty \cos x^2 dx$, and $\int_0^\infty \sin x^2 dx$* , American Mathematical Monthly **97** (1990), 39–42.
- [255] S. Weintraub, *Early transcendentals*, American Mathematical Monthly **104** (1997), 623–631.
- [256] S. Weintraub, *An observation on average velocity*, American Mathematical Monthly **130** (2023), 384.

- [257] G. Weiss, *Complex methods in harmonic analysis*, American Mathematical Monthly **77** (1970), 465–474.
- [258] D. Wend, *Uniqueness of solutions of ordinary differential equations*, American Mathematical Monthly **74** (1967), 948–950.
- [259] D. Widder, *What is the Laplace transform?*, American Mathematical Monthly **52** (1945), 410–425.
- [260] D. Widder, *Advanced Calculus*, 3rd edition, Dover, 1989.
- [261] D. Willett, *The existence-uniqueness theorem for an n th order linear ordinary differential equation*, American Mathematical Monthly **75** (1968), 174–178.
- [262] A. Wouk, *On the Cauchy–Picard method*, American Mathematical Monthly **70** (1963), 158–162.
- [263] B. Youse, *Introduction to Real Analysis*, Allyn and Bacon, 1972.
- [264] J. van Yzeren, *Classroom notes: A rehabilitation of $(1 + z/n)^n$* , American Mathematical Monthly **77** (1970), 995–998.
- [265] J. van Yzeren, *Classroom notes: Moivre’s and Fresnel’s integrals by simple integration*, American Mathematical Monthly **86** (1979), 690–693.
- [266] H. Zatzkis, *Volume and surface of a sphere in N -dimensional Euclidean space*, Mathematics Magazine **30** (1957), 155–157.
- [267] A. Ziebur, *Classroom notes: Uniqueness theorems for ordinary differential equations*, American Mathematical Monthly **69** (1962), 421–423.

Index

- Abel summability, 118
- absolute convergence, 29, 285
- absolute value, 2, 5
- $\mathcal{B}(X, Y)$, 19
- $\mathcal{B}_N(X, \mathbf{C}^m)$, 302
- $\mathcal{B}_N(X, \mathbf{R}^m)$, 302
- Baire category theorem, 81, 84
- Banach spaces, 281
- bases, 70
- bilipschitz mappings, 147
- bounded convergence theorem, 280
- bounded linear functionals, 289
- bounded linear mappings, 286
- bounded mappings, 19
- bounded sequences, 17
- bounded sets, 16
- Brouwer's fixed-point theorem, 153
- \mathbf{C} , 5
- $c(X, \mathbf{C})$, 56
- $c(X, \mathbf{C}^m)$, 302
- $c(X, \mathbf{R})$, 56
- $c(X, \mathbf{R}^m)$, 302
- $C(X, Y)$, 23
- $(C, 1)$ convergence of sequences, 121
- $(C, 1)$ summable series, 121
- \mathbf{C}^n , 5
- $C_0(X, \mathbf{C})$, 310
- $c_0(X, \mathbf{C})$, 36
- $C_0(X, \mathbf{R})$, 310
- $c_0(X, \mathbf{R})$, 36
- $c_{00}(X, \mathbf{C})$, 9
- $c_{00}(X, \mathbf{R})$, 9
- $C_b(X, Y)$, 23
- $C_{com}(X, \mathbf{C})$, 307
- $C_{com}(X, \mathbf{R})$, 307
- $C_N(X, \mathbf{C}^m)$, 303
- $C_N(X, \mathbf{R}^m)$, 303
- $C_{b,N}(X, \mathbf{C}^m)$, 303
- $C_{b,N}(X, \mathbf{R}^m)$, 303
- Cartesian products, 207
- Cauchy products, 113
- Cauchy sequences, 12
- Cauchy subsequence property, 77
- Cauchy–Schwarz inequality, 5, 26, 33, 41, 47, 53, 273, 284
- cells, 304
- Cesaro means, 121
- closed balls, 17
- compact support, 307
- complementary linear subspaces, 200
- complete metric spaces, 12
- complex conjugate, 117
- complex exponential function, 117
- complex linear mappings, 133
- connected sets, 176
- continuous differentiability, 163
- contraction mapping theorem, 152
- convergent series, 28, 284
- convex sets, 299
- Cramer's rule, 194
- dense sets, 54
- derivative of a function, 2
- determinants, 141, 193
- diameter of a set, 65
- differentiable mappings, 172
- differentiable mappings, 2, 166
- differential of a mapping, 167
- dimensions of vector spaces, 198
- directional derivatives, 164
- discrete metric, 306
- dominated convergence theorem, 268

- dual norms, 289
- dual spaces, 288
- eigenvalues, 143
- eigenvectors, 143
- equicontinuity, 88
- equiconvergence of limits, 102
- Euclidean metric, 4, 6
- Euclidean norm, 4, 5
- Euler's identity, 118
- extended real numbers, 245
- extreme value theorem, 241
- first category, 84
- fixed points, 151
- generalized convergence, 260
- $GL(\mathbf{R}^n)$, 184
- gradient of a function, 167
- graphs of mappings, 237
- Hahn–Banach theorem, 290
- Hilbert spaces, 284
- Hilbert–Schmidt norm, 133, 138
- Hölder continuity, 293
- homeomorphisms, 241
- identity matrix, 140
- identity operator, 140
- implicit function theorem, 192
- inner products, 126, 127, 129–131, 277, 278, 282, 283
- interior of a set, 83
- inverse function theorem, 186
- invertible linear mappings, 184
- isometries, 22, 145, 146
- iterated integrals, 220
- joint continuity, 216
- kernels of linear mappings, 199
- $\mathcal{L}(\mathbf{C}^n, \mathbf{C}^m)$, 137
- $\mathcal{L}(\mathbf{R}^n)$, 184
- $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$, 174
- $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$, 137
- ℓ^1 , 30
- $\ell^1(X, \mathbf{C})$, 253
- $\ell^1(X, \mathbf{R})$, 253
- $\ell^1(\mathbf{Z}_+, \mathbf{C})$, 30
- $\ell^1(\mathbf{Z}_+, \mathbf{R})$, 30
- ℓ^2 , 32
- $\ell^2(X, \mathbf{C})$, 274
- $\ell^2(X, \mathbf{R})$, 274
- $\ell^2(\mathbf{Z}_+, \mathbf{C})$, 32
- $\ell^2(\mathbf{Z}_+, \mathbf{R})$, 32
- ℓ^∞ , 23
- $\ell^\infty(X, \mathbf{C})$, 22
- $\ell^\infty(X, \mathbf{R})$, 22
- $\ell_N^\infty(X, \mathbf{C}^m)$, 302
- $\ell_N^\infty(X, \mathbf{R}^m)$, 302
- Lebesgue integrals, 26, 279
- limit point property, 72
- Lindelöf property, 72
- Lindelöf's theorem, 71
- linear functionals, 288
- linear mappings, 133
- linear subspaces, 56, 198, 282
- $\text{Lip}(f)$, 143
- $\text{Lip}(X, Y)$, 143
- Lipschitz mappings, 1, 143
 - of order α , 293
- local compactness, 308
- locally constant mappings, 177
- locally Lipschitz mappings, 179
- lower semicontinuity, 240
- matrices, 132
- meager sets, 84
- Minkowski's inequality, 7
- monotone convergence theorem, 267
- nonmeager sets, 84
- norms, 3, 5, 9, 23, 25, 39, 49, 132, 137, 281
- nowhere dense sets, 84
- nowhere vanishing, 57
- null spaces of linear mappings, 199
- open balls, 15
- open mappings, 156
- operator norms, 138, 287
- orthogonal vectors, 286

- partial derivatives, 162
- partial Lipschitz conditions, 161, 215
- partitions of intervals, 42, 221
- partitions of unity, 230
- path-connected sets, 301
- Peano's theorem, 96
- piecewise-linear functions, 221
- pointwise boundedness, 97
- pointwise convergence, 13
- pointwise Lipschitz conditions, 168
- projections, 200
- \mathbf{R} , 2
- \mathbf{R}^n , 3
- radius of convergence, 108
- rank of a linear mapping, 196
- rank theorem, 198, 202
- rearrangements, 112
- Riemann–Stieltjes integrals, 42
- Riesz representation theorem, 289
- second category, 84
- semicontinuity, 240
- seminorms, 44, 52, 299
- separable metric spaces, 69
- separate continuity, 216
- separated sets, 176
- separates points, 57
- sequential compactness, 75
- small subsequence property, 77
- square-summable functions, 272
- Stone–Weierstrass theorem, 55, 59
- subalgebras, 56
- summable functions, 250, 253
- sums of linear subspaces, 198
- supports of functions, 9, 49, 306
- supremum metric, 20
- supremum norm, 23, 40, 49, 303
- totally bounded sets, 67
- transverse linear subspaces, 199
- triangle inequality, 3
- uniform boundedness, 97
- uniform closure, 60
- uniform continuity, 216
- uniform convergence, 13
- uniform equicontinuity, 89
- uniformly Cauchy sequences, 90
- uniformly closed sets of functions, 60
- upper semicontinuity, 240
- V' , 289
- V^* , 289
- V^{alg} , 288
- vanishing at infinity, 36, 309
- Weierstrass' criterion, 106
- Weierstrass' theorem, 55
- \mathbf{Z}_+ , 9