

THE CALCULUS INTEGRAL

(2009)

Brian S. Thomson
Simon Fraser University



CLASSICALREALANALYSIS.COM

This text is intended for a rigorous course introducing integration theory to calculus students, but starting at a level preceding the more rigorous courses in real analysis. Any student with a suitably thorough course on derivatives should be able to handle the first few chapters of the integration theory without trouble. Since “all” exercises are worked through in the appendix, the text is particularly well suited to self-study.

For further information on this title and others in the series visit our website.

www.classicalrealanalysis.com

There are PDF files of all of our texts freely available for download as well as instructions on how to order trade paperback copies.

Cover Image: Sir Isaac Newton

And from my pillow, looking forth by light
Of moon or favouring stars, I could behold
The antechapel where the statue stood
Of Newton with his prism and silent face,
The marble index of a mind for ever
Voyaging through strange seas of Thought, alone.

... William Wordsworth, *The Prelude*.

Citation: *The Calculus Integral*, Brian S. Thomson, ClassicalRealAnalysis.com (2010), xiv xxx pp. [ISBN 1442180951]

Date PDF file compiled: July 19, 2009

BETA VERSION β 0.2

The file or paperback that you are reading should be considered a work in progress. In a classroom setting make sure all participants are using the same beta version. We will add and amend, depending on feedback from our users, until the text appears to be in a stable condition.

ISBN: 1442180951

EAN-13: 9781442180956



CLASSICALREALANALYSIS.COM

PREFACE

There are plenty of calculus books available, many free or at least cheap, that discuss integrals. Why add another one?

Our purpose is to present integration theory at a calculus level and in an easier manner by defining the definite integral in a very traditional way, but a way that avoids the equally traditional Riemann sums definition.

Riemann sums enter the picture, to be sure, but the integral is defined in the way that Newton himself would surely endorse. Thus the fundamental theorem of the calculus starts off as the definition and the relation with Riemann sums becomes a theorem (not the definition of the definite integral as has, most unfortunately, been the case for many years).

As usual in mathematical presentations we all end up in the same place. It is just that we have taken a different route to get there. It is only a pedagogical issue of which route offers the clearest perspective. The common route of starting with the definition of the Riemann integral, providing the then necessary detour into improper integrals, and ultimately heading towards the Lebesgue integral is arguably not the best path although it has at least the merit of historical fidelity.

Acknowledgments

I have used without comment material that has appeared in the textbook

[TBB] *Elementary Real Analysis*, 2nd Edition, B. S. Thomson, J. B. Bruckner, A. M. Bruckner, Classical-RealAnalysis.com (2008).

I wish to express my thanks to my co-authors for permission to recycle that material into the idiosyncratic form that appears here and their encouragement (or at least lack of discouragement) in this project.

I would also like to thank the following individuals who have offered feedback on the material, or who have supplied interesting exercises or solutions to our exercises: *[your name here]*, ...

Note to the instructor

Since it is possible that some brave mathematicians will undertake to present integration theory to undergraduates students using the presentation in this text, it would be appropriate for us to address some comments to them.

What should I teach the weak calculus students?

Let me dispense with this question first. Don't teach them this material. I also wouldn't teach them the Riemann integral. I think a reasonable outline for these students would be this:

- (1). An informal account of the indefinite integral formula

$$\int F'(x) dx = F(x) + C$$

just as an antiderivative notation with a justification provided by the mean-value theorem.

- (2). An account of what it means for a function to be continuous on an interval $[a, b]$.

- (3). The definition

$$\int_a^b F'(x) dx = F(b) - F(a)$$

for continuous functions $F : [a, b] \rightarrow \mathbb{R}$ that are differentiable at all but finitely many points in (a, b) . The mean-value theorem again justifies the definition. You won't need improper integrals, e.g., because of (3),

$$\int_0^1 \frac{1}{\sqrt{x}} dx = \int_0^1 \frac{d}{dx} (2\sqrt{x}) dx = 2 - 0.$$

- (4). Any properties of integrals that are direct translations of derivative properties.

(5). The Riemann sums *identity*

$$\int_a^b f(x) dx = \sum_{i=1}^n f(\xi_i^*)(x_i - x_{i-1})$$

where the points ξ_i^* that make this precise are selected (yet again) by the mean-value theorem.

(6). The Riemann sums *approximation*

$$\int_a^b f(x) dx \approx \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1})$$

where the points ξ_i can be freely selected inside the interval. Continuity of f justifies this since $f(\xi_i) \approx f(\xi_i^*)$.

That's all! On the other hand, for students that are not considered marginal, the presentation in the text should lead to a full theory of integration on the real line provided at first that the student is sophisticated enough to handle ϵ , δ arguments and simple compactness proofs (notably Bolzano-Weierstrass and Cousin lemma proofs).

Why the calculus integral?

Perhaps the correct question is “Why not the Lebesgue integral?” After all, integration theory on the real line is not adequately described by either the calculus integral or the Riemann integral.

The answer that we all seem to have agreed upon is that Lebesgue's theory is too difficult for beginning students of integration theory. Thus we need a “teaching integral,” one that will present all the usual rudiments of the theory in way that prepares the student for the later introduction of measure and integration.

Using the Riemann integral as a teaching integral requires starting with summations and a difficult and awkward limit formulation. Eventually one reaches the fundamental theorem of the calculus. The fastest and most efficient way of teaching integration theory on the real line is, instead, at the outset to interpret the calculus integral

$$\int_a^b F'(x) dx = F(b) - F(a)$$

as a definition. The primary tool is the very familiar mean-value theorem. That theorem leads quickly back to Riemann sums in any case.

The instructor must then drop the habit of calling this the fundamental theorem of the calculus. Within a few lectures the main properties of integrals are available and all of the computational exercises are accessible. This is because everything is merely an immediate application of differentiation theorems. There is no need for an “improper” theory of

the integral since integration of unbounded functions requires no additional ideas or lectures.

There is a long and distinguished historical precedent for this kind of definition. For all of the 18th century the integral was understood *only* in this sense¹ The descriptive definition of the Lebesgue integral, which too can be taken as a starting point, is exactly the same: but now requires F to be absolutely continuous and F' is defined only almost everywhere. The Denjoy-Perron integral has the same descriptive definition but relaxes the condition on F to that of generalized absolute continuity. Thus the narrative of integration theory on the real line can told simply as an interpretation of the integral as meaning merely

$$\int_a^b F'(x) dx = F(b) - F(a).$$

Why not the Riemann integral?

Or you may prefer to persist in teaching to your calculus students the Riemann integral and its ugly step-sister, the improper Riemann integral. There are many reasons for ceasing to use this as a teaching integral; the web page, “*Top ten reasons for dumping the Riemann integral*” which you can find on our site

www.classicalrealanalysis.com

has a tongue-in-cheek account of some of these.

The Riemann integral does not do a particularly good job of introducing integration theory to students. That is not to say that students should be sheltered from the notion of Riemann sums. It is just that a whole course confined to the Riemann integral wastes considerable time on a topic and on methods that are not worthy of such devotion.

In this presentation the Riemann sums approximations to integrals simply enters into the discussion naturally by way of the mean-value theorem of the differential calculus. It does not require several lectures on approximations of areas and other motivating stories.

¹Certainly Newton and his followers saw it in this sense. For Leibnitz and his advocates the integral was a sum of infinitesimals, but that only explained the connection with the derivative. For a lucid account of the thinking of the mathematicians to whom we owe all this theory see Judith V. Grabiner, *Who gave you the epsilon? Cauchy and the origins of rigorous calculus*, American Mathematical Monthly 90 (3), 1983, 185–194.

The calculus integral

For all of the 18th century and a good bit of the 19th century integration theory, as we understand it, was simply the subject of antidifferentiation. Thus what we would call the fundamental theorem of the calculus would have been considered a tautology: that is how an integral is defined. Both the differential and integral calculus are, then, simply, the study of derivatives with the integral calculus largely focussed on the inverse problem.

This is often expressed by modern analysts by claiming that the *Newton integral* of a function $f : [a, b] \rightarrow \mathbb{R}$ is defined as

$$\int_a^b f(x) dx = F(b) - F(a)$$

where $F : [a, b] \rightarrow \mathbb{R}$ is any continuous function whose derivative $F'(x)$ is identical with $f(x)$ at all points $a < x < b$. While Newton would have used no such notation or terminology, he would doubtless agree with us that this is precisely the integral he intended.

The technical justification for this definition of the Newton integral is little more than the mean-value theorem of the calculus. Thus it is ideally suited for teaching integration theory to beginning students of the calculus. Indeed, it would be a reasonable bet that most students of the calculus drift eventually into a hazy world of little-remembered lectures and eventually think that this is exactly what an integral is anyway. Certainly it is the only method that they have used to compute integrals.

For these reasons we have called it the *calculus integral*². But none of us teach the calculus integral. Instead we teach the Riemann integral. Then, when the necessity of integrating unbounded functions arise, we teach the improper Riemann integral. When the student is more advanced we sheepishly let them know that the integration theory that they have learned is just a moldy 19th century concept that was replaced in all serious studies a full century ago. We do not apologize for the fact that we have misled them; indeed we likely will not even mention the fact that the improper Riemann integral and the Lebesgue integral are quite distinct; most students accept the mantra that the Lebesgue integral is better and they take it for granted that it includes what they learned. We also do not point out just how awkward and misleading the Riemann theory is: we just drop the subject entirely.

Why is the Riemann integral the “teaching integral” of choice when the calculus integral offers a better and easier approach to integration theory? The transition from the Riemann integral to the Lebesgue integral requires abandoning Riemann sums in favor of measure theory. The transition from the improper Riemann integral to the Lebesgue integral

²The play on the usual term “integral calculus” is intentional.

is usually flubbed.

The transition from the calculus integral to the Lebesgue integral (and beyond) can be made quite logically. Introduce, first, sets of measure zero and some simple related concepts. Then an integral which completely includes the calculus integral and yet is as general as one requires can be obtained by repeating Newton's definition above: the *integral* of a function $f : [a, b] \rightarrow \mathbb{R}$ is defined as

$$\int_a^b f(x) dx = F(b) - F(a)$$

where $F : [a, b] \rightarrow \mathbb{R}$ is any continuous function whose derivative $F'(x)$ is identical with $f(x)$ at all points $a < x < b$ with the exception of a set of points N that is of measure zero and on which F has zero variation.

We are employing here the usual conjuror's trick that mathematicians often use. We take some late characterization of a concept and reverse the presentation by taking that as a definition. One will see all the familiar theory gets presented along the way but that, because the order is turned on its head, quite a different perspective emerges.

Give it a try and see if it works for your students. By the end of the Part One the student will have learned the calculus integral, seen all of the familiar integration theorems of the integral calculus, worked with Riemann sums, functions of bounded variation, studied countable sets and sets of measure zero, and given a working definition of the Lebesgue integral.

Part Two returns to the general calculus integral and gives the full Henstock-Kurzweil characterization. Chapter 6 presents the theory of the Lebesgue measure and integral on the line and all of the usual material of a first graduate course, but again in an untraditional manner. By this point the student is ready for a typical graduate course in abstract measure theory. If you choose to present only Part One (the elementary calculus integral) your students should still be as adequately prepared for their studies as the usual route through the Riemann integral would have done. Maybe *better* prepared.

Contents

Preface	i
Note to the instructor	iii
Table of Contents	viii
 I Elementary Theory of the Integral	 xxvii
1 What you should know first	1
1.1 What is the calculus about?	2
1.2 What is an interval?	2
1.3 Sequences and series	5
1.3.1 Sequences	5
1.3.2 Series	8
1.4 Partitions	9
1.4.1 Cousin's partitioning argument	10
1.5 What is a function?	11
1.6 Continuous functions	12
1.6.1 Uniformly continuous and continuous functions	13

1.6.2	Oscillation of a function	17
1.6.3	Endpoint limits	18
1.6.4	Boundedness properties	22
1.7	Existence of maximum and minimum	24
1.7.1	The Darboux property of continuous functions	25
1.8	Derivatives	27
1.9	Differentiation rules	30
1.10	Mean-value theorem	30
1.10.1	Rolle's theorem	31
1.10.2	Mean-Value theorem	32
1.10.3	The Darboux property of the derivative	37
1.10.4	Vanishing derivatives and constant functions	38
1.10.5	Vanishing derivatives with exceptional sets	38
1.11	Lipschitz functions	40
2	The Indefinite Integral	43
2.1	An indefinite integral on an interval	44
2.1.1	Role of the finite exceptional set	45
2.1.2	Features of the indefinite integral	46
2.1.3	The notation $\int f(x) dx$	46
2.2	Existence of indefinite integrals	48
2.2.1	Upper functions	49
2.2.2	The main existence theorem	50
2.3	Basic properties of indefinite integrals	51
2.3.1	Linear combinations	52
2.3.2	Integration by parts	52
2.3.3	Change of variable	53
2.3.4	What is the derivative of the indefinite integral?	54
2.3.5	Partial fractions	55
2.3.6	Tables of integrals	58

3	The Definite Integral	59
3.1	Definition of the calculus integral	60
3.1.1	Alternative definition of the integral	61
3.1.2	Infinite integrals	62
3.1.3	Simple properties of integrals	63
3.1.4	Integrability of bounded functions	66
3.1.5	Integrability for the unbounded case	67
3.1.6	Products of integrable functions	69
3.1.7	Notation: $\int_a^a f(x) dx$ and $\int_b^a f(x) dx$	70
3.1.8	The dummy variable: what is the “ x ” in $\int_a^b f(x) dx$?	70
3.1.9	Definite vs. indefinite integrals	71
3.1.10	The calculus student’s notation	72
3.2	Mean-value theorems for integrals	73
3.3	Riemann sums	75
3.3.1	Exact computation by Riemann sums	76
3.3.2	Uniform Approximation by Riemann sums	79
3.3.3	Theorem of G. A. Bliss	81
3.3.4	Pointwise approximation by Riemann sums	83
3.4	Properties of the integral	85
3.4.1	Inequalities	85
3.4.2	Linear combinations	86
3.4.3	Subintervals	86
3.4.4	Integration by parts	87
3.4.5	Change of variable	87
3.4.6	What is the derivative of the definite integral?	89
3.5	Absolute integrability	90
3.5.1	Functions of bounded variation	91
3.5.2	Indefinite integrals and bounded variation	95
3.6	Sequences and series of integrals	95
3.6.1	The counterexamples	96

3.6.2	Uniform convergence	102
3.6.3	Uniform convergence and integrals	108
3.6.4	A defect of the calculus integral	110
3.6.5	Uniform limits of continuous derivatives	111
3.6.6	Uniform limits of discontinuous derivatives	113
3.7	The monotone convergence theorem	114
3.7.1	Summing inside the integral	114
3.7.2	Monotone convergence theorem	116
3.8	Integration of power series	117
3.9	Applications of the integral	126
3.9.1	Area and the method of exhaustion	128
3.9.2	Volume	131
3.9.3	Length of a curve	134
3.10	Numerical methods	136
3.10.1	Maple methods	142
3.10.2	Maple and infinite integrals	144
3.11	More Exercises	145
4	Beyond the calculus integral	147
4.1	Countable sets	148
4.1.1	Cantor's theorem	149
4.2	Derivatives which vanish outside of countable sets	150
4.2.1	Calculus integral [countable set version]	150
4.3	Sets of measure zero	152
4.3.1	The Cantor dust	154
4.4	The Devil's staircase	158
4.4.1	Construction of Cantor's function	158
4.5	Functions with zero variation	161
4.5.1	Zero variation lemma	163
4.5.2	Zero derivatives imply zero variation	164

CONTENTS

xiii

4.5.3	Continuity and zero variation	164
4.5.4	Absolute continuity	165
4.5.5	Absolute continuity in Vitali's sense	167
4.6	The integral	168
4.6.1	Infinite integrals	171
4.7	Lipschitz functions and bounded integrable functions	172
4.8	Approximation by Riemann sums	173
4.9	Properties of the integral	174
4.9.1	Inequalities	174
4.9.2	Linear combinations	175
4.9.3	Subintervals	175
4.9.4	Integration by parts	176
4.9.5	Change of variable	176
4.9.6	What is the derivative of the definite integral?	177
4.9.7	Monotone convergence theorem	177
4.9.8	Summation of series theorem	178
4.9.9	Null functions	179
4.10	The Henstock-Kurweil integral	180
4.11	The Lebesgue integral	181
4.12	The Riemann integral	183

II Theory of the Integral on the Real Line 185

5	Covering Theorems	187
5.1	Covering Relations	188
5.1.1	Partitions and subpartitions	188
5.1.2	Covering relations	189
5.1.3	Prunings	189
5.1.4	Full covers	190

5.1.5	Fine covers	190
5.1.6	Uniformly full covers	190
5.1.7	Cousin covering lemma	194
5.1.8	Decomposition of full covers	195
5.1.9	Riemann sums	196
5.2	Sets of measure zero	200
5.2.1	Lebesgue measure of open sets	200
5.2.2	Sets of Lebesgue measure zero	202
5.2.3	Sequences of measure zero sets	202
5.2.4	Almost everywhere language	206
5.3	Full null sets	207
5.4	Fine null sets	209
5.5	The Mini-Vitali Covering Theorem	210
5.5.1	Covering lemmas for families of compact intervals	211
5.5.2	Proof of the Mini-Vitali covering theorem	212
5.6	Functions having zero variation	215
5.6.1	Zero variation and zero derivatives	217
5.6.2	Generalization of the zero derivative/variation	218
5.6.3	Absolutely continuous functions	220
5.6.4	Absolute continuity and derivatives	221
5.7	Lebesgue differentiation theorem	223
5.7.1	Upper and lower derivates	224
5.7.2	Geometrical lemmas	225
5.7.3	Proof of the Lebesgue differentiation theorem	226
6	The Integral	231
6.1	The integral and integrable functions	232
6.1.1	Infinite integrals	233
6.1.2	Approximation by Riemann sums	234
6.2	The Henstock-Kurzweil characterization of the integral	237

6.2.1	Definition of Henstock and Kurzweil	237
6.2.2	Upper and lower integrals	238
6.2.3	The integral and integrable functions	240
6.2.4	First Cauchy criterion	242
6.2.5	Second Cauchy criterion	243
6.2.6	Proof of equivalence	245
6.3	Elementary properties of the integral	250
6.3.1	Integration and order	250
6.3.2	Integration of linear combinations	251
6.3.3	Integrability on subintervals	251
6.3.4	Additivity	251
6.3.5	Change of variable	251
6.3.6	Integration by parts	252
6.3.7	Derivative of the integral	253
6.3.8	Null functions	254
6.3.9	Monotone convergence theorem	255
6.3.10	Summing inside the integral	256
6.3.11	Two convergence lemmas	257
6.4	Equi-integrability	262
7	Lebesgue's Integral	263
7.1	The Lebesgue integral	264
7.2	Lebesgue measure	264
7.2.1	Basic property of Lebesgue measure	265
7.3	Vitali covering theorem	266
7.3.1	Classical version of Vitali's theorem	267
7.3.2	Proof that $\lambda = \lambda^* = \lambda_*$	269
7.4	Density theorem	270
7.5	Additivity	272
7.6	Measurable sets	275

7.6.1	Definition of measurable sets	275
7.6.2	Properties of measurable sets	275
7.6.3	Increasing sequences of sets	277
7.6.4	Existence of nonmeasurable sets	278
7.7	Measurable functions	280
7.7.1	Continuous functions are measurable	280
7.7.2	Derivatives and integrable functions are measurable	281
7.7.3	Simple functions	282
7.7.4	Series of simple functions	283
7.7.5	Limits of measurable functions	284
7.8	Construction of the integral	285
7.8.1	Characteristic functions of measurable sets	285
7.8.2	Characterizations of measurable sets	286
7.8.3	Integral of simple functions	288
7.8.4	Integral of nonnegative measurable functions	288
7.8.5	Fatou's Lemma	289
7.8.6	Derivatives of functions of bounded variation	292
7.8.7	Characterization of the Lebesgue integral	293
7.8.8	McShane's Criterion	294
7.8.9	Nonabsolutely integrable functions	298
7.9	The Lebesgue integral as a set function	299
7.10	Characterizations of the indefinite integral	303
7.10.1	Integral of nonnegative, integrable functions	305
7.10.2	Integral of absolutely integrable functions	305
7.10.3	Integral of nonabsolutely integrable functions	306
7.10.4	Proofs	306
7.11	Denjoy's program	309
7.12	The Riemann integral	310

8 Stieltjes Integrals

313

8.1	Stieltjes integrals	313
8.1.1	Definition of the Stieltjes integral	315
8.1.2	Henstock's zero variation criterion	318
8.2	Regulated functions	319
8.3	Variation expressed as an integral	322
8.4	Representation theorems for functions of bounded variation	324
8.4.1	Jordan decomposition	324
8.4.2	Jordan decomposition theorem: differentiation	325
8.4.3	Representation by saltus functions	327
8.4.4	Representation by singular functions	327
8.5	Reducing a Stieltjes integral to an ordinary integral	327
8.6	Properties of the indefinite integral	330
8.6.1	Existence of the integral from derivative statements	334
8.7	Existence of the Stieltjes integral for continuous functions	335
8.8	Integration by parts	336
8.9	Lebesgue-Stieltjes measure	338
8.10	Mutually singular functions	341
8.11	Singular functions	343
8.12	Length of curves	344
8.12.1	Formula for the length of curves	345
9	Nonabsolutely Integrable Functions	349
9.1	Variational Measures	350
9.1.1	Full and fine variational measures	351
9.1.2	Finite variation and σ -finite variation	352
9.1.3	The Vitali property	353
9.1.4	Kolmogorov equivalence	353
9.1.5	Variation of continuous, increasing functions	354
9.1.6	Variation and image measure	355
9.1.7	Variational classifications of real functions	356

9.2	Derivates and variation	359
9.2.1	Ordinary derivates and variation	359
9.2.2	Dini derivatives and variation	360
9.2.3	Lipschitz numbers	362
9.2.4	Six growth lemmas	364
9.3	Continuous functions with σ -finite variation	369
9.3.1	Variation on compact sets	370
9.3.2	λ -absolutely continuous functions	372
9.4	Vitali property and differentiability	372
9.5	The Vitali property and variation	374
9.5.1	Monotonic functions	374
9.5.2	Functions of bounded variation	375
9.5.3	Functions of σ -finite variation	375
9.6	Characterization of the Vitali property	376
9.7	Characterization of λ -absolute continuity	377
9.8	Mapping properties	378
9.9	Lusin's conditions	380
9.10	Banach-Zarecki Theorem	380
9.11	Local Lebesgue integrability conditions	382
9.12	Continuity of upper and lower integrals	386
9.13	A characterization of the integral	387
9.14	Integral of Dini derivatives	391
9.14.1	Motivation	391
9.14.2	Quasi-Cousin covering lemma	393
9.14.3	Estimates of integrals from derivates	394
9.14.4	Estimates of integrals from Dini derivatives	395
10	Integration in $\mathbb{R}^n$	399
10.1	Some background	399
10.1.1	Intervals and covering relations	400

CONTENTS

xix

10.2	Measure and integral	402
10.2.1	Lebesgue measure in $\mathbb{R}^n$	403
10.2.2	The fundamental lemma	403
10.3	Measurable sets and measurable functions	406
10.3.1	Measurable functions	407
10.3.2	Notation	409
10.4	General measure theory	410
10.5	Iterated integrals	411
10.5.1	Formulation of the iterated integral property	413
10.5.2	Fubini's theorem	416
10.6	Expression as a Stieltjes integral	417
11	Appendix	419
11.1	Glossary of terms	419
11.1.1	absolute continuity	419
11.1.2	absolute convergence	420
11.1.3	absolute convergence test	421
11.1.4	absolute integration	421
11.1.5	almost everywhere	421
11.1.6	Baire category theorem	422
11.1.7	Bolzano-Weierstrass argument	424
11.1.8	bounded set	424
11.1.9	bounded function	425
11.1.10	bounded sequence	425
11.1.11	bounded variation	425
11.1.12	bounded monotone sequence argument	425
11.1.13	Cantor dust	426
11.1.14	Cauchy sequences	426
11.1.15	characteristic function of a set	427
11.1.16	closed set	427

11.1.17 compactness argument	427
11.1.18 connected set	428
11.1.19 convergence of a sequence	428
11.1.20 component of an open set	428
11.1.21 composition of functions	429
11.1.22 constant of integration	429
11.1.23 continuous function	429
11.1.24 contraposition	430
11.1.25 converse	430
11.1.26 countable set	431
11.1.27 Cousin's partitioning argument	431
11.1.28 Cousin's covering argument	432
11.1.29 Darboux property	433
11.1.30 definite integral	433
11.1.31 De Morgan's Laws	433
11.1.32 dense	434
11.1.33 derivative	434
11.1.34 Devil's staircase	434
11.1.35 domain of a function	435
11.1.36 empty set	435
11.1.37 equivalence relation	435
11.1.38 graph of a function	435
11.1.39 partition	436
11.1.40 Henstock-Kurzweil integral	436
11.1.41 indefinite integral	437
11.1.42 indirect proof	437
11.1.43 infs and sups	438
11.1.44 integers	438
11.1.45 integral test for series	438
11.1.46 intermediate value property	439

11.1.47 interval	439
11.1.48 least upper bound argument	439
11.1.49 induction	440
11.1.50 inverse of a function	440
11.1.51 isolated point	441
11.1.52 Jordan decomposition	441
11.1.53 Lebesgue integral	441
11.1.54 limit of a function	442
11.1.55 linear combination	444
11.1.56 Lipschitz function	444
11.1.57 locally bounded function	444
11.1.58 lower bound of a set	444
11.1.59 managing epsilons	445
11.1.60 meager	445
11.1.61 mean-value theorem	446
11.1.62 measure zero	446
11.1.63 monotone subsequence argument	447
11.1.64 mostly everywhere	447
11.1.65 natural numbers	448
11.1.66 nearly everywhere	448
11.1.67 negations of quantified statements	448
11.1.68 nested interval argument	449
11.1.69 nowhere dense	449
11.1.70 open set	450
11.1.71 one-to-one and onto function	450
11.1.72 ordered pairs	450
11.1.73 oscillation of a function	451
11.1.74 partition	451
11.1.75 perfect set	452
11.1.76 pointwise continuous function	452

11.1.77 preimage of a function	453
11.1.78 quantifiers	453
11.1.79 range of a function	454
11.1.80 rational numbers	454
11.1.81 real numbers	454
11.1.82 relations	455
11.1.83 residual	455
11.1.84 Riemann sum	455
11.1.85 Riemann integral	456
11.1.86 series	457
11.1.87 set-builder notation	458
11.1.88 set notation	458
11.1.89 subpartition	459
11.1.90 summation by parts	459
11.1.91 sups and infs	459
11.1.92 subsets, unions, intersection, and differences	460
11.1.93 total variation function	460
11.1.94 uniformly continuous function	460
11.1.95 upper bound of a set	461
11.1.96 variation of a function	461
11.2 Answers to exercises	463

Part I

Elementary Theory of the Integral

Chapter 1

What you should know first

This chapter begins a review of the differential calculus. We go, perhaps, deeper than the reader has gone before because we need to justify and prove everything we shall do. If your calculus courses so far have left the proofs of certain theorems (most notably the existence of maxima and minima of continuous functions) to a “more advanced course” then this will be, indeed, deeper. If your courses proved such theorems then there is nothing here in Chapters 1–3 that is essentially harder.

The text is about the integral calculus. The entire theory of integration can be presented as an attempt to solve the equation

$$\frac{dy}{dx} = f(x)$$

for a suitable function $y = F(x)$. Certainly we cannot approach such a problem until we have some considerable expertise in the study of derivatives. So that is where we begin. Well-informed, or smug students, may skip over this chapter and begin immediately with the integration theory. The indefinite integral starts in Chapter 2. The definite integral continues in Chapter 3. The material in Chapter 4 takes the integration theory, which up to this point has been at an elementary level, to the next stage.

We assume the reader knows the rudiments of the calculus and can answer the majority of the exercises here without much trouble. Later chapters will introduce topics in a very careful order. Here we assume in advance that you know basic facts about functions, limits, continuity, derivatives, sequences and series and need only a careful review.

1.1 What is the calculus about?

The calculus is the study of the derivative and the integral. In fact, the integral is so closely related to the derivative that the study of the integral is an essential part of studying derivatives. Thus there is really one topic only: the derivative. Most university courses are divided, however, into the separate topics of Differential Calculus and Integral Calculus, to use the old-fashioned names.

Your main objective in studying the calculus is to understand (thoroughly) what the concepts of derivative and integral are and to comprehend the many relations among the concepts.

It may seem to a typical calculus student that the subject is mostly all about computations and algebraic manipulations. While that may appear to be the main feature of the courses it is, by no means, the main objective.

If you can remember yourself as a child learning arithmetic perhaps you can put this in the right perspective. A child's point of view on the study of arithmetic centers on remembering the numbers, memorizing addition and multiplication tables, and performing feats of mental arithmetic. The goal is actually, though, what some people have called numeracy: familiarity and proficiency in the world of numbers. We all know that the computations themselves can be trivially performed on a calculator and that the mental arithmetic skills of the early grades are not an end in themselves.

You should think the same way about your calculus problems. In the end you need to understand what all these ideas mean and what the structure of the subject is. Ultimately you are seeking mathematical literacy, the ability to think in terms of the concepts of the calculus. In your later life you will most certainly not be called upon to differentiate a polynomial or integrate a trigonometric expression (unless you end up as a drudge teaching calculus to others). But, if we are successful in our teaching of the subject, you will be able to understand and use many of the concepts of economics, finance, biology, physics, statistics, etc. that are expressible in the language of derivatives and integrals.

1.2 What is an interval?

We should really begin with a discussion of the real numbers themselves, but that would add a level of complexity to the text that is not completely necessary. If you need a full treatment of the real numbers see our text [TBB]¹. Make sure especially to understand the use of suprema and infima in working with real numbers. We can begin by defining what

¹Thomson, Bruckner, Bruckner, *Elementary Real Analysis*, 2nd Edition (2008). The relevant chapters are available for free download at classicalrealanalysis.com.

1.2. WHAT IS AN INTERVAL?

we mean by those sets of real numbers called *intervals*.

All of the functions of the elementary calculus are defined on intervals or on sets that are unions of intervals. This language, while simple, should be clear.

An interval is the collection of all the points on the real line that lie between two given points [the endpoints], or the collection of all points that lie on the right or left side of some point. The endpoints are included for closed intervals and not included for open intervals.

Here is the notation and language: Take any real numbers a and b with $a < b$. Then the following symbols describe *intervals* on the real line:

- **(open bounded interval)** (a, b) is the set of all real numbers between (but not including) the points a and b , i.e., all $x \in \mathbb{R}$ for which $a < x < b$.
- **(closed, bounded interval)** $[a, b]$ is the set of all real numbers between (and including) the points a and b , i.e., all $x \in \mathbb{R}$ for which $a \leq x \leq b$.
- **(half-open bounded interval)** $[a, b)$ is the set of all real numbers between (but not including b) the points a and b , i.e., all $x \in \mathbb{R}$ for which $a \leq x < b$.
- **(half-open bounded interval)** $(a, b]$ is the set of all real numbers between (but not including a) the points a and b , i.e., all $x \in \mathbb{R}$ for which $a < x \leq b$.
- **(open unbounded interval)** (a, ∞) is the set of all real numbers greater than (but not including) the point a , i.e., all $x \in \mathbb{R}$ for which $a < x$.
- **(open unbounded interval)** $(-\infty, b)$ is the set of all real numbers lesser than (but not including) the point b , i.e., all $x \in \mathbb{R}$ for which $x < b$.
- **(closed unbounded interval)** $[a, \infty)$ is the set of all real numbers greater than (and including) the point a , i.e., all $x \in \mathbb{R}$ for which $a \leq x$.
- **(closed unbounded interval)** $(-\infty, b]$ is the set of all real numbers lesser than (and including) the point b , i.e., all $x \in \mathbb{R}$ for which $x \leq b$.

- **(the entire real line)** $(-\infty, \infty)$ is the set of all real numbers. This can be reasonably written as all x for which $-\infty < x < \infty$.

Exercise 1 Do the symbols $-\infty$ and ∞ stand for real numbers? What are they then? Answer ☐

Exercise 2 (bounded sets) Which intervals are bounded? [To find out what a bounded set is see page 424.] Answer ☐

Exercise 3 (open sets) Show that an open interval (a, b) or (a, ∞) or $(-\infty, b)$ is an open set. [To find out what an open set is see page 450.] Answer ☐

Exercise 4 (closed sets) Show that a closed interval $[a, b]$ or $[a, \infty)$ or $(-\infty, b]$ is a closed set. [To find out what a closed set is see page 427.] Answer ☐

Exercise 5 Show that the intervals $[a, b)$ and $(a, b]$ are neither closed nor open. Answer ☐

Exercise 6 (intersection of two open intervals) Is the intersection of two open intervals an open interval? Answer ☐

Exercise 7 (intersection of two closed intervals) Is the intersection of two closed intervals a closed interval? Answer ☐

Exercise 8 Is the intersection of two unbounded intervals an unbounded interval? Answer ☐

Exercise 9 When is the union of two open intervals an open interval? Answer ☐

Exercise 10 When is the union of two closed intervals an open interval? Answer ☐

Exercise 11 Is the union of two bounded intervals a bounded set? Answer ☐

Exercise 12 If I is an open interval and C is a finite set what kind of set might be $I \setminus C$? Answer ☐

Exercise 13 If I is a closed interval and C is a finite set what kind of set might be $I \setminus C$? Answer ☐

1.3. SEQUENCES AND SERIES

1.3 Sequences and series

We will need the method of sequences and series in our studies of the integral. In this section we present a brief review. By a *sequence* we mean an infinite list of real numbers

$$s_1, s_2, s_3, s_4, \dots$$

and by a *series* we mean that we intend to sum the terms in some sequence

$$a_1 + a_2 + a_3 + a_4 + \dots$$

The notation for such a sequence would be $\{s_n\}$ and for such a series $\sum_{k=1}^{\infty} a_k$.

1.3.1 Sequences

A sequence converges to a number L if the terms of the sequence eventually get close to (and remain close to) the number L . A sequence is Cauchy if the terms of the sequence eventually get close together (and remain close together). The notions are very intimately related.

Definition 1.1 (convergent sequence) A sequence of real numbers $\{s_n\}$ is said to converge to a real number L if, for every $\varepsilon > 0$ there is an integer N so that

$$L - \varepsilon < s_n < L + \varepsilon$$

for all integers $n \geq N$. In that case we write

$$\lim_{n \rightarrow \infty} s_n = L.$$

If a sequence fails to converge it is said to diverge.

Definition 1.2 (Cauchy sequence) A sequence of real numbers $\{s_n\}$ is said to be a Cauchy sequence if, for every $\varepsilon > 0$ there is an integer N so that

$$|s_n - s_m| < \varepsilon$$

for all pairs of integers $n, m \geq N$.

Definition 1.3 (divergent to ∞) A sequence of real numbers $\{s_n\}$ is said to diverge to ∞ if, for every real number M there is an integer N so that $s_n > M$ for all integers $n \geq N$. In that case we write

$$\lim_{n \rightarrow \infty} s_n = \infty.$$

[We do not say the sequence “converges to ∞ .”]

In the exercises you will show that every convergent sequence is a Cauchy sequence and, conversely, that every Cauchy sequence is a convergent sequence. We will also need to review the behavior of monotone sequences and of subsequences. All of the exercises should be looked at as the techniques discussed here are used freely throughout the rest of the material of the text.

Exercise 14 A sequence $\{s_n\}$ is said to be bounded if there is a number M so that $|s_n| \leq M$ for all n . Show that every convergent sequence is bounded. Give an example of a bounded sequence that is not convergent. Answer $\square$

Exercise 15 Show that every Cauchy sequence is bounded. Give an example of a bounded sequence that is not Cauchy. Answer $\square$

Exercise 16 Show that every convergent sequence is Cauchy. [The converse is proved below after we have looked for convergent subsequences.] Answer $\square$

Exercise 17 (theory of sequence limits) Suppose that $\{s_n\}$ and $\{t_n\}$ are convergent sequences.

1. What can you say about the sequence $x_n = as_n + bt_n$ for real numbers a and b ?
2. What can you say about the sequence $y_n = s_nt_n$?
3. What can you say about the sequence $y_n = \frac{s_n}{t_n}$?
4. What can you say if $s_n \leq t_n$ for all n ?

Answer $\square$

1.3. SEQUENCES AND SERIES

Exercise 18 A sequence $\{s_n\}$ is said to be nondecreasing [or monotone nondecreasing] if

$$s_1 \leq s_2 \leq s_3 \leq s_4 \leq \dots$$

Show that such a sequence is convergent if and only if it is bounded, and in fact that

$$\lim_{n \rightarrow \infty} s_n = \sup\{s_n : n = 1, 2, 3, \dots\}.$$

Answer ☐

Exercise 19 (nested interval argument) A sequence $\{[a_n, b_n]\}$ of closed, bounded intervals is said to be a nested sequence of intervals shrinking to a point if

$$[a_1, b_1] \supset [a_2, b_2] \supset [a_3, b_3] \supset [a_4, b_4] \supset \dots$$

and

$$\lim_{n \rightarrow \infty} (b_n - a_n) = 0.$$

Show that there is a unique point in all of the intervals.

Answer ☐

Exercise 20 Given a sequence $\{s_n\}$ and a sequence of integers

$$1 \leq n_1 < n_2 < n_3 < n_4 < \dots$$

construct the new sequence

$$\{s_{n_k}\} = s_{n_1}, s_{n_2}, s_{n_3}, s_{n_4}, s_{n_5}, \dots$$

The new sequence is said to be a subsequence of the original sequence. Show that every sequence $\{s_n\}$ has a subsequence that is monotone, i.e., either monotone nondecreasing

$$s_{n_1} \leq s_{n_2} \leq s_{n_3} \leq s_{n_4} \leq \dots$$

or else monotone nonincreasing

$$s_{n_1} \geq s_{n_2} \geq s_{n_3} \geq s_{n_4} \geq \dots$$

Answer ☐

Exercise 21 (Bolzano-Weierstrass property) Show that every bounded sequence has a convergent subsequence.

Answer ☐

Exercise 22 Show that every Cauchy sequence is convergent. [The converse was proved earlier.] Answer ☐

Exercise 23 Let E be a closed set and $\{x_n\}$ a convergent sequence of points in E . Show that $x = \lim_{n \rightarrow \infty} x_n$ must also belong to E . Answer ☐

1.3.2 Series

The theory of series reduces to the theory of sequence limits by interpreting the sum of the series to be the sequence limit

$$\sum_{k=1}^{\infty} a_k = \lim_{n \rightarrow \infty} \sum_{k=1}^n a_k.$$

Definition 1.4 (convergent series) A series

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + a_4 + \dots$$

is said to be convergent and to have a sum equal to L if the sequence of partial sums

$$S_n = \sum_{k=1}^n a_k = a_1 + a_2 + a_3 + a_4 + \dots + a_n$$

converges to the number L . If a series fails to converge it is said to diverge.

Definition 1.5 (absolutely convergent series) A series

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + a_4 + \dots$$

is said to be absolutely convergent if both of the sequences of partial sums

$$S_n = \sum_{k=1}^n a_k = a_1 + a_2 + a_3 + a_4 + \dots + a_n$$

and

$$T_n = \sum_{k=1}^n |a_k| = |a_1| + |a_2| + |a_3| + |a_4| + \dots + |a_n|$$

are convergent.

1.4. PARTITIONS

Exercise 24 *Let*

$$S_n = \sum_{k=1}^n a_k = a_1 + a_2 + a_3 + a_4 + \cdots + a_n$$

be the sequence of partial sums of a series

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + a_4 + \dots$$

Show that S_n is Cauchy if and only if for every $\varepsilon > 0$ there is an integer N so that

$$\left| \sum_{k=m}^n a_k \right| < \varepsilon$$

*for all $n \geq m \geq N$.*Answer $\square$ **Exercise 25** *Let*

$$S_n = \sum_{k=1}^n a_k = a_1 + a_2 + a_3 + a_4 + \cdots + a_n$$

and

$$T_n = \sum_{k=1}^n |a_k| = |a_1| + |a_2| + |a_3| + |a_4| + \cdots + |a_n|.$$

Show that if $\{T_n\}$ is a Cauchy sequence then so too is the sequence $\{S_n\}$. What can you conclude from this? Answer $\square$ **1.4 Partitions**

When working with an interval and functions defined on intervals we shall frequently find that we must subdivide the interval at a finite number of points. For example if $[a, b]$ is a closed, bounded interval then any finite selection of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

breaks the interval into a collection of subintervals

$$\{[x_{i-1}, x_i] : i = 1, 2, 3, \dots, n\}$$

that are nonoverlapping and whose union is all of the original interval $[a, b]$.

Most often when we do this we would need to focus attention on certain points chosen from each of the intervals. If ξ_i is a point in $[x_{i-1}, x_i]$ then the collection

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

will be called a *partition* of the interval $[a, b]$.

In sequel we shall see many occasions when splitting up an interval this way is useful. In fact our integration theory for a function f defined on the interval $[a, b]$ can often be expressed by considering the sum

$$\sum_{k=1}^n f(\xi_k)(x_k - x_{k-1})$$

over a partition. This is known as a *Riemann sum* for f .

1.4.1 Cousin's partitioning argument

The simple lemma we need for many proofs was first formulated by Pierre Cousin.

Lemma 1.6 (Cousin) *For every point x in a closed, bounded interval $[a, b]$ let there be given a positive number $\delta(x)$. Then there must exist at least one partition*

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of the interval $[a, b]$ with the property that each interval $[x_{i-1}, x_i]$ has length smaller than $\delta(\xi_i)$.

Exercise 26 Show that this lemma is particularly easy if $\delta(x) = \delta$ is constant for all x in $[a, b]$.

Answer ☐

Exercise 27 Prove Cousin's lemma using a nested interval argument.

Answer ☐

Exercise 28 Prove Cousin's lemma using a "last point" argument.

Answer ☐

Exercise 29 Use Cousin's lemma to prove this version of the Heine-Borel theorem: Let C be a collection of open intervals covering a closed, bounded interval $[a, b]$. Then there is a finite subcollection $\{(c_i, d_i) : i = 1, 2, 3, \dots, n\}$ from C that also covers $[a, b]$.

Answer ☐

Exercise 30 (connected sets) A set of real numbers E is disconnected if it is possible to find two disjoint open sets G_1 and G_2 so that both sets contain at least one point of E and together they include all of E . Otherwise a set is connected. Show that the interval $[a, b]$ is connected using a Cousin partitioning argument. Answer $\square$

Exercise 31 (connected sets) Show that the interval $[a, b]$ is connected using a last point argument. Answer $\square$

Exercise 32 Show that a set E that contains at least two points is connected if and only if it is an interval. Answer $\square$

1.5 What is a function?

For most calculus students a function is a formula. We use the symbol

$$f : E \rightarrow \mathbb{R}$$

to indicate a function (whose name is “ f ”) that must be defined at every point x in the set E (E must be, for this course, a subset of $\mathbb{R}$) and to which some real number value $f(x)$ is assigned. The way in which $f(x)$ is assigned need not, of course, be some algebraic formula. Any method of assignment is possible as long as it is clear what is the *domain* of the function [i.e., the set E] and what is the value [i.e., $f(x)$] that this function assumes at each point x in E .

More important is the concept itself. When we see

“Let $f : [0, 1] \rightarrow \mathbb{R}$ be the function defined by $f(x) = x^2$ for all x in the interval $[0, 1]$...”

or just simply

“Let $g : [0, 1] \rightarrow \mathbb{R}$...”

we should be equally comfortable. In the former case we know and can compute every value of the function f and we can sketch its graph. In the latter case we are just asked to consider that some function g is under consideration: we know that it has a value $g(x)$ at every point in its domain (i.e., the interval $[0, 1]$) and we know that it has a graph and we can discuss that function g as freely as we can the function f .

Even so calculus students will spend, unfortunately for their future understanding, undue time with formulas. For this remember one rule: if a function is specified by a formula it is also essential to know what is the domain of the

function. The convention is usually to specify exactly what the domain intended should be, or else to take the largest possible domain that the formula given would permit. Thus $f(x) = \sqrt{x}$ does not specify a function until we reveal what the domain of the function should be; since $f(x) = \sqrt{x}$ ($0 \leq x < \infty$) is the best we could do, we would normally claim that the domain is $[0, \infty)$.

Exercise 33 In a calculus course what are the assumed domains of the trigonometric functions $\sin x$, $\cos x$, and $\tan x$?

Answer ☐

Exercise 34 In a calculus course what are the assumed domains of the inverse trigonometric functions $\arcsin x$ and $\arctan x$?

Answer ☐

Exercise 35 In a calculus course what are the assumed domains of the exponential and natural logarithm functions e^x and $\log x$?

Answer ☐

Exercise 36 In a calculus course what might be the assumed domains of the functions given by the formulas

$$f(x) = \frac{1}{(x^2 - x - 1)^2}, \quad g(x) = \sqrt{x^2 - x - 1}, \quad \text{and} \quad h(x) = \arcsin(x^2 - x - 1)?$$

Answer ☐

1.6 Continuous functions

Most of the functions that one encounters in the calculus are continuous. Continuity refers to the idea that a function f should have small increments $f(d) - f(c)$ on small intervals $[c, d]$. That is, however, a horribly imprecise statement of it; what we wish is that the increment $f(d) - f(c)$ should be as small as we please provided that the interval $[c, d]$ is sufficiently small.

The interpretation of

... as small as ... provided ... is sufficiently small ...

is invariably expressed in the language of ϵ , δ , definitions that you will encounter in all of your mathematical studies and which it is essential to master. Nearly everything in this course is expressed in ϵ , δ language.

1.6.1 Uniformly continuous and continuous functions

The notion of uniform continuity below is a global condition: it is a condition which holds throughout the whole of some interval. Often we will encounter a more local variant where the continuity condition holds only close to some particular point in the interval where the function is defined. We fix a particular point x_0 in the interval and then repeat the definition of uniform continuity but with the extra requirement that it need hold only near the point x_0 .

Definition 1.7 (uniform continuity) Let $f : I \rightarrow \mathbb{R}$ be a function defined on an interval I . We say that f is uniformly continuous if for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$|f(d) - f(c)| < \varepsilon$$

whenever c, d are points in I for which $|d - c| < \delta$.

Definition 1.8 (pointwise continuity) Let $f : I \rightarrow \mathbb{R}$ be a function defined on an open interval I and let x_0 be a point in that interval. We say that f is [pointwise] continuous at x_0 if for every $\varepsilon > 0$ there is a $\delta(x_0) > 0$ so that

$$|f(x) - f(x_0)| < \varepsilon$$

whenever x is a point in I for which $|x - x_0| < \delta(x_0)$. We say f is continuous on the open interval I provided f is continuous at each point of I .

Note that continuity at a point requires that the function is defined on both sides of the point as well as at the point. Thus we would be very cautious about asserting continuity of the function $f(x) = \sqrt{x}$ at 0. Uniform continuity on an interval $[a, b]$ does not require that the function is defined on the right of a or the left of b . We are comfortable asserting that $f(x) = \sqrt{x}$ is uniformly continuous on $[0, 1]$. (It is.)

A comment on the language: For most textbooks the language is simply

“continuous on a set” vs. “uniformly continuous on a set”

and the word “pointwise” is dropped. For teaching purposes it is important to grasp the distinction between these two definitions; we use here the pointwise/uniform language to emphasize this very important distinction. We will see

this same idea and similar language in other places. A sequence of functions can converge *pointwise* or *uniformly*. A Riemann sum approximation to an integral can be *pointwise* or *uniform*.

Exercise 37 Show that uniform continuity is stronger than pointwise continuity, i.e., show that a function $f(x)$ that is uniformly continuous on an open interval I is necessarily continuous on that interval. Answer ☐

Exercise 38 Show that uniform continuity is strictly stronger than pointwise continuity, i.e., show that a function $f(x)$ that is continuous on an open interval I is not necessarily uniformly continuous on that interval. Answer ☐

Exercise 39 Construct a function that is defined on the interval $(-1, 1)$ and is continuous only at the point $x_0 = 0$. Answer ☐

Exercise 40 Show that the function $f(x) = x$ is uniformly continuous on the interval $(-\infty, \infty)$. Answer ☐

Exercise 41 Show that the function $f(x) = x^2$ is not uniformly continuous on the interval $(-\infty, \infty)$. Answer ☐

Exercise 42 Show that the function $f(x) = x^2$ is uniformly continuous on any bounded interval. Answer ☐

Exercise 43 Show that the function $f(x) = x^2$ is not uniformly continuous on the interval $(-\infty, \infty)$ but is continuous at every real number x_0 . Answer ☐

Exercise 44 Show that the function $f(x) = \frac{1}{x}$ is not uniformly continuous on the interval $(0, \infty)$ or on the interval $(-\infty, 0)$ but is continuous at every real number $x_0 \neq 0$. Answer ☐

Exercise 45 (linear combinations) Suppose that F and G are functions on an open interval I and that both of them are continuous at a point x_0 in that interval. Show that any linear combination $H(x) = rF(x) + sG(x)$ must also be continuous at the point x_0 . Does the same statement apply to uniform continuity? Answer ☐

Exercise 46 (products) Suppose that F and G are functions on an open interval I and that both of them are continuous at a point x_0 in that interval. Show that the product $H(x) = F(x)G(x)$ must also be continuous at the point x_0 . Does the same statement apply to uniform continuity? Answer ☐

1.6. CONTINUOUS FUNCTIONS

15

Exercise 47 (quotients) Suppose that F and G are functions on an open interval I and that both of them are continuous at a point x_0 in that interval. Must the quotient $H(x) = F(x)/G(x)$ must also be pointwise continuous at the point x_0 . Is there a version for uniform continuity? Answer ☐

Exercise 48 (compositions) Suppose that F is a function on an open interval I and that F is continuous at a point x_0 in that interval. Suppose that every value of F is contained in an interval J . Now suppose that G is a function on the interval J that is continuous at the point $z_0 = f(x_0)$. Show that the composition function $H(x) = G(F(x))$ must also be continuous at the point x_0 . Answer ☐

Exercise 49 Show that the absolute value function $f(x) = |x|$ is uniformly continuous on every interval. ☐

Exercise 50 Show that the function

$$D(x) = \begin{cases} 1 & \text{if } x \text{ is irrational,} \\ \frac{1}{n} & \text{if } x = \frac{m}{n} \text{ in lowest terms,} \end{cases}$$

where m, n are integers expressing the rational number $x = \frac{m}{n}$, is continuous at every irrational number but discontinuous at every rational number. ☐

Exercise 51 (Heaviside's function) Step functions play an important role in integration theory. They offer a crude way of approximating functions. The function

$$H(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

is a simple step function that assumes just two values, 0 and 1, where 0 is assumed on the interval $(-\infty, 0)$ and 1 is assumed on $[0, \infty)$. Find all points of continuity of H . Answer ☐

Exercise 52 (step Functions) A function f defined on a bounded interval is a step function if it assumes finitely many values, say $b_1, b_2, \dots, b_N$ and for each $1 \leq i \leq N$ the set

$$f^{-1}(b_i) = \{x : f(x) = b_i\},$$

which represents the set of points at which f assumes the value b_i , is a finite union of intervals and singleton point sets. (See Figure 1.1 for an illustration.) Find all points of continuity of a step function. Answer ☐

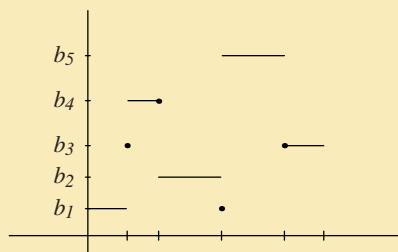


Figure 1.1: Graph of a step function.

Exercise 53 (characteristic function of the rationals) Show that function defined by the formula

$$R(x) = \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} |\cos(m! \pi x)|^n$$

is discontinuous at every point.

Answer ☐

Exercise 54 (distance of a closed set to a point) Let C be a closed set and define a function by writing

$$d(x, C) = \inf\{|x - y| : y \in C\}.$$

This function gives a meaning to the distance between a set C and a point x . If $x_0 \in C$, then $d(x_0, C) = 0$, and if $x_0 \notin C$, then $d(x_0, C) > 0$. Show that function is continuous at every point. How might you interpret the fact that the distance function is continuous?

Answer ☐

Exercise 55 (sequence definition of continuity) Prove that a function f defined on an open interval is continuous at a point x_0 if and only if $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$ for every sequence $\{x_n\} \rightarrow x_0$.

Answer ☐

Exercise 56 (mapping definition of continuity) Let $f : (a, b) \rightarrow \mathbb{R}$ be defined on an open interval. Then f is continuous on (a, b) if and only if for every open set $V \subset \mathbb{R}$, the set

$$f^{-1}(V) = \{x \in A : f(x) \in V\}$$

is open.

Answer ☐

1.6.2 Oscillation of a function

Continuity of a function f asserts that the increment of f on an interval (c, d) , i.e., the value $f(d) - f(c)$, must be small if the interval $[c, d]$ is small. This can often be expressed more conveniently by the oscillation of the function on the interval $[c, d]$.

Definition 1.9 Let f be a function defined on an interval I . We write

$$\omega f(I) = \sup\{|f(x) - f(y)| : x, y \in I\}$$

and call this the oscillation of the function f on the interval I .

Exercise 57 Establish these properties of the oscillation:

1. $\omega f([c, d]) \leq \omega f([a, b])$ if $[c, d] \subset [a, b]$.
2. $\omega f([a, c]) \leq \omega f([a, b]) + \omega f([b, c])$ if $a < b < c$.

□

Exercise 58 (uniform continuity and oscillations) Let $f : I \rightarrow \mathbb{R}$ be a function defined on an interval I . Show that f is uniformly continuous on I if and only if, for every $\varepsilon > 0$, there is a $\delta > 0$ so that

$$\omega f([c, d]) < \varepsilon$$

whenever $[c, d]$ is a subinterval of I for which $|d - c| < \delta$.

[Thus uniformly continuous functions have small increments $f(d) - f(c)$ or equivalently small oscillations $\omega f([c, d])$ on sufficiently small intervals.]

Answer □

Exercise 59 (uniform continuity and oscillations) Show that f is a uniformly continuous function on a closed, bounded interval $[a, b]$ if and only if, for every $\varepsilon > 0$, there are points

$$a = x_0 < x_1 < x_2 < x_3 < \cdots < x_{n-1} < x_n = b$$

so that each of

$$\omega f([x_0, x_1]), \omega f([x_1, x_2]), \dots, \text{ and } \omega f([x_{n-1}, x_n])$$

is smaller than ε . (Is there a similar statement for uniform continuity on open intervals?)

Answer □

Exercise 60 (continuity and oscillations) Show that f is continuous at a point x_0 in an open interval I if and only if for every $\varepsilon > 0$ there is a $\delta(x_0) > 0$ so that

$$\omega f([x_0 - \delta(x_0), x_0 + \delta(x_0)]) \leq \varepsilon.$$

Answer $\square$

Exercise 61 (continuity and oscillations) Let $f : I \rightarrow \mathbb{R}$ be a function defined on an open interval I . Show that f is continuous at a point x_0 in I if and only if for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$\omega f([c, d]) < \varepsilon$$

whenever $[c, d]$ is a subinterval of I that contains the point x_0 and for which $|d - c| < \delta$.

Answer $\square$

Exercise 62 (limits and oscillations) Suppose that f is defined on a bounded open interval (a, b) . Show that a necessary and sufficient condition in order that $F(a+) = \lim_{x \rightarrow a+} F(x)$ should exist is that for all $\varepsilon > 0$ there should exist a positive number $\delta(a)$ so that

$$\omega f((a, a + \delta(a))) < \varepsilon.$$

Answer $\square$

Exercise 63 (infinite limits and oscillations) Suppose that F is defined on (∞, ∞) . Show that a necessary and sufficient condition in order that $F(\infty) = \lim_{x \rightarrow \infty} F(x)$ should exist is that for all $\varepsilon > 0$ there should exist a positive number T so that

$$\omega f((T, \infty)) < \varepsilon.$$

Show that the same statement is true for $F(-\infty) = \lim_{x \rightarrow -\infty} F(x)$ with the requirement that

$$\omega f((-\infty, -T)) < \varepsilon.$$

Answer $\square$

1.6.3 Endpoint limits

We are interested in computing, if possible the one-sided limits

$$F(a+) = \lim_{x \rightarrow a+} F(x) \text{ and } F(b-) = \lim_{x \rightarrow b-} F(x)$$

for a function defined on a bounded, open interval (a, b) . Pointwise continuity will not help us here, but uniform continuity does and this is an important application of uniform continuity.

There is a close connection between uniform and pointwise continuity. Uniform continuity is the stronger condition. But under the conditions stated in our next theorem continuity and uniform continuity are equivalent. This theorem should be attributed to Cauchy but cannot be, for he failed to notice the difference between the two concepts and simply took it for granted that they were equivalent.

Theorem 1.10 (endpoint limits) *Let $F : (a, b) \rightarrow \mathbb{R}$ be a function that is continuous on the bounded, open interval (a, b) . Then the two limits*

$$F(a+) = \lim_{x \rightarrow a+} F(x) \text{ and } F(b-) = \lim_{x \rightarrow b-} F(x)$$

exist if and only if F is uniformly continuous on (a, b) .

Corollary 1.11 (extension property) *Let $F : (a, b) \rightarrow \mathbb{R}$ be a function that is continuous on the bounded, open interval (a, b) . Then F can be extended to a uniformly continuous function on all of the closed, bounded interval $[a, b]$ if and only if F is uniformly continuous on (a, b) . That extension is obtained by defining*

$$F(a) = F(a+) = \lim_{x \rightarrow a+} F(x) \text{ and } F(b) = F(b-) = \lim_{x \rightarrow b-} F(x)$$

both of which limits exist if F is uniformly continuous on (a, b) .

Corollary 1.12 (subinterval property) *Let $F : (a, b) \rightarrow \mathbb{R}$ be a function that is continuous on the bounded, open interval (a, b) . Then F is uniformly continuous on every closed, bounded subinterval $[c, d] \subset (a, b)$, but may or may not be a uniformly continuous function on all of (a, b) .*

Corollary 1.13 (monotone property) *Let $F : (a, b) \rightarrow \mathbb{R}$ be a function that is continuous on the bounded, open interval (a, b) and is either monotone nondecreasing or monotone nonincreasing. Then F is uniformly continuous on (a, b) if and only if F is bounded on (a, b) .*

Exercise 64 Prove one direction of the endpoint limit theorem [Theorem 1.10]: Show that if F is uniformly continuous on (a, b) then the two limits

$$F(a+) = \lim_{x \rightarrow a+} F(x) \text{ and } F(b-) = \lim_{x \rightarrow b-} F(x)$$

exist.

Answer ☐

Exercise 65 Prove the other direction of the endpoint limit theorem [Theorem 1.10] using Exercise 62 and a Cousin partitioning argument: Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is continuous on the bounded, open interval (a, b) and that the two limits

$$F(a+) = \lim_{x \rightarrow a+} F(x) \text{ and } F(b-) = \lim_{x \rightarrow b-} F(x)$$

exist. Show that F is uniformly continuous on (a, b) .

Answer ☐

Exercise 66 Prove the extension property [Corollary 1.11].

Answer ☐

Exercise 67 Prove the subinterval property [Corollary 1.12].

Answer ☐

Exercise 68 Prove the monotone property [Corollary 1.13].

Answer ☐

Exercise 69 Prove the other direction of the endpoint limit theorem using a Bolzano-Weierstrass compactness argument: Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is continuous on the bounded, open interval (a, b) and that the two limits

$$F(a+) = \lim_{x \rightarrow a+} F(x) \text{ and } F(b-) = \lim_{x \rightarrow b-} F(x)$$

exist. Show that F is uniformly continuous on (a, b) .

Answer ☐

Exercise 70 Prove the other direction of the endpoint limit theorem using a Heine-Borel argument: Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is continuous on the bounded, open interval (a, b) and that the two limits

$$F(a+) = \lim_{x \rightarrow a+} F(x) \text{ and } F(b-) = \lim_{x \rightarrow b-} F(x)$$

exist. Show that F is uniformly continuous on (a, b) .

Answer ☐

1.6. CONTINUOUS FUNCTIONS

21

Exercise 71 Show that the theorem fails if we drop the requirement that the interval is bounded.

Answer ☐

Exercise 72 Show that the theorem fails if we drop the requirement that the interval is closed.

Answer ☐

Exercise 73 Criticize this proof of the false theorem that if f is continuous on an interval (a, b) then f must be uniformly continuous on (a, b) .

Suppose if f is continuous on (a, b) . Let $\epsilon > 0$ and for any x_0 in (a, b) choose a $\delta > 0$ so that $|f(x) - f(x_0)| < \epsilon$ if $|x - x_0| < \delta$. Then if c and d are any points that satisfy $|c - d| < \delta$ just set $c = x$ and $d = x_0$ to get $|f(d) - f(c)| < \epsilon$. Thus f must be uniformly continuous on (a, b) .

Answer ☐

Exercise 74 Suppose that $G : (a, b) \rightarrow \mathbb{R}$ is continuous at every point of an open interval (a, b) . Then show that G is uniformly continuous on every closed, bounded subinterval $[c, d] \subset (a, b)$.

Answer ☐

Exercise 75 Show that, if $F : (a, b) \rightarrow \mathbb{R}$ is a function that is continuous on the bounded, open interval (a, b) but not uniformly continuous, then one of the two limits

$$F(a+) = \lim_{x \rightarrow a+} F(x) \text{ or } F(b-) = \lim_{x \rightarrow b-} F(x)$$

must fail to exist.

Answer ☐

Exercise 76 Show that, if $F : (a, b) \rightarrow \mathbb{R}$ is a function that is continuous on the bounded, open interval (a, b) and both of the two limits

$$F(a+) = \lim_{x \rightarrow a+} F(x) \text{ and } F(b-) = \lim_{x \rightarrow b-} F(x)$$

exist then F is in fact uniformly continuous on (a, b) .

Answer ☐

Exercise 77 Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is a function defined on an open interval (a, b) and that c is a point in that interval. Show that F is continuous at c if and only if both of the two one-sided limits

$$F(c+) = \lim_{x \rightarrow c+} F(x) \text{ and } F(c-) = \lim_{x \rightarrow c-} F(x)$$

exist and $F(c) = F(c+) = F(c-)$.

Answer ☐

1.6.4 Boundedness properties

Continuity has boundedness implications. Pointwise continuity supplies local boundedness; uniform continuity supplies global boundedness, but only on bounded intervals.

Definition 1.14 (bounded function) Let $f : I \rightarrow \mathbb{R}$ be a function defined on an interval I . We say that f is bounded on I if there is a number M so that

$$|f(x)| \leq M$$

for all x in the interval I .

Definition 1.15 (locally bounded function) A function f defined on an interval I is said to be locally bounded at a point x_0 if there is a $\delta(x_0) > 0$ so that f is bounded on the set

$$(x_0 - \delta(x_0), x_0 + \delta(x_0)) \cap I.$$

Theorem 1.16 Let $f : I \rightarrow \mathbb{R}$ be a function defined on a bounded interval I and suppose that f is uniformly continuous on I . Then f is a bounded function on I .

Theorem 1.17 Let $f : I \rightarrow \mathbb{R}$ be a function defined on an open interval I and suppose that f is continuous at a point x_0 in I . Then f is locally bounded at x_0 .

Remember that, if f is continuous on an open interval (a, b) , then f is uniformly continuous on each closed subinterval $[c, d] \subset (a, b)$. Thus, in order for f to be unbounded on (a, b) the large values are occurring only at the endpoints. Let us say that f is locally bounded on the right at a if there is at least one interval $(a, a + \delta_a)$ on which f is bounded. Similarly we can define locally bounded on the left at b . This corollary is then immediate.

Corollary 1.18 Let $f : (a, b) \rightarrow \mathbb{R}$ be a function defined on an open interval (a, b) . Suppose that

1. f is continuous at every point in (a, b) .
2. f is locally bounded on the right at a .
3. f is locally bounded on the left at b .

Then f is bounded on the interval (a, b) .

Exercise 78 Use Exercise 59 to prove Theorem 1.16.

Answer ☐

Exercise 79 Prove Theorem 1.17 by proving that all continuous functions are locally bounded.

Answer ☐

Exercise 80 It follows from Theorem 1.16 that a continuous, unbounded function on a bounded open interval (a, b) cannot be uniformly continuous. Can you prove that a continuous, bounded function on a bounded open interval (a, b) must be uniformly continuous?

Answer ☐

Exercise 81 Show that f is not bounded on an interval I if and only if there must exist a sequence of points $\{x_n\}$ for which $f(x_n) \rightarrow \infty$.

Answer ☐

Exercise 82 Using Exercise 81 and the Bolzano-Weierstrass argument, show that if a function f is locally bounded at each point of a closed, bounded interval $[a, b]$ then f must be bounded on $[a, b]$. ☐

Exercise 83 Using Cousin's lemma, show that if a function f is locally bounded at each point of a closed, bounded interval $[a, b]$ then f must be bounded on $[a, b]$. ☐

Exercise 84 If a function is uniformly continuous on an unbounded interval must the function be unbounded? Could it be bounded? ☐

Answer ☐

Exercise 85 Suppose $f, g : I \rightarrow \mathbb{R}$ are two bounded functions on I . Is the sum function $f + g$ necessarily bounded on I ? Is the product function fg necessarily bounded on I ? ☐

Answer ☐

Exercise 86 Suppose $f, g : I \rightarrow \mathbb{R}$ are two bounded functions on I and suppose that the function g does not assume the value zero. Is the quotient function f/g necessarily bounded on I ? Answer ☐

Exercise 87 Suppose $f, g : \mathbb{R} \rightarrow \mathbb{R}$ are two bounded functions. Is the composite function $h(x) = f(g(x))$ necessarily bounded? Answer ☐

Exercise 88 Show that the function $f(x) = \sin x$ is uniformly continuous on the interval $(-\infty, \infty)$. Answer ☐

Exercise 89 A function defined on an interval I is said to satisfy a Lipschitz condition there if there is a number M with the property that

$$|F(x) - F(y)| \leq M|x - y|$$

for all $x, y \in I$. Show that a function that satisfies a Lipschitz condition on an interval is uniformly continuous on that interval. Answer ☐

Exercise 90 Show that f is not uniformly continuous on an interval I if and only if there must exist two sequences of points $\{x_n\}$ and $\{y_n\}$ from that interval for which $x_n - y_n \rightarrow 0$ but $f(x_n) - f(y_n)$ does not converge to zero. Answer ☐

1.7 Existence of maximum and minimum

Uniformly continuous functions are bounded on bounded intervals. Must they have a maximum and a minimum value? We know that continuous functions need not be bounded so our focus will be on uniformly continuous functions on closed, bounded intervals.

Theorem 1.19 Let $F : [a, b] \rightarrow \mathbb{R}$ be a function defined on a closed, bounded interval $[a, b]$ and suppose that F is uniformly continuous on $[a, b]$. Then F attains both a maximum value and a minimum value in that interval.

Exercise 91 Prove Theorem 1.19 using a least upper bound argument. Answer ☐

1.7. EXISTENCE OF MAXIMUM AND MINIMUM

25

Exercise 92 Prove Theorem 1.19 using a Bolzano-Weierstrass argument.

Answer ☐

Exercise 93 Give an example of a uniformly continuous function on the interval $(0, 1)$ that attains a maximum but does not attain a minimum.

Answer ☐

Exercise 94 Give an example of a uniformly continuous function on the interval $(0, 1)$ that attains a minimum but does not attain a maximum.

Answer ☐

Exercise 95 Give an example of a uniformly continuous function on the interval $(0, 1)$ that attains neither a minimum nor a maximum.

Answer ☐

Exercise 96 Give an example of a uniformly continuous function on the interval $(-\infty, \infty)$ that attains neither a minimum nor a maximum.

Answer ☐

Exercise 97 Give an example of a uniformly continuous, bounded function on the interval $(-\infty, \infty)$ that attains neither a minimum nor a maximum.

Answer ☐

Exercise 98 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be an everywhere continuous function with the property that

$$\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow -\infty} f(x) = 0.$$

Show that f has either an absolute maximum or an absolute minimum but not necessarily both.

Answer ☐

Exercise 99 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be an everywhere continuous function that is periodic in the sense that for some number p , $f(x + p) = f(x)$ for all $x \in \mathbb{R}$. Show that f has an absolute maximum and an absolute minimum.

Answer ☐

1.7.1 The Darboux property of continuous functions

We define the Darboux property of a function and show that all continuous functions have this property.

Definition 1.20 (Darboux Property) Let f be defined on an interval I . Suppose that for each $a, b \in I$ with $f(a) \neq f(b)$, and for each d between $f(a)$ and $f(b)$, there exists c between a and b for which $f(c) = d$. We then say that f has the Darboux property [intermediate value property] on I .

Functions with this property are called *Darboux* functions after Jean Gaston Darboux (1842–1917), who showed in 1875 that for every differentiable function F on an interval I , the derivative F' has the intermediate value property on I .

Theorem 1.21 (Darboux property of continuous functions) Let $f : (a, b) \rightarrow \mathbb{R}$ be a continuous function on an open interval (a, b) . Then f has the Darboux property on that interval.

Exercise 100 Prove Theorem 1.21 using a Cousin covering argument.

Answer ☐

Exercise 101 Prove Theorem 1.21 using a Bolzano-Weierstrass argument.

Answer ☐

Exercise 102 Prove Theorem 1.21 using the Heine-Borel property.

Answer ☐

Exercise 103 Prove Theorem 1.21 using the least upper bound property.

Answer ☐

Exercise 104 Suppose that $f : (a, b) \rightarrow \mathbb{R}$ is a continuous function on an open interval (a, b) . Show that f maps (a, b) onto an interval. Show that this interval need not be open, need not be closed, and need not be bounded.

Answer ☐

Exercise 105 Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is a uniformly continuous function on a closed, bounded interval $[a, b]$. Show that f maps $[a, b]$ onto an interval. Show that this interval must be closed and bounded.

Answer ☐

Exercise 106 Define the function

$$F(x) = \begin{cases} \sin x^{-1} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Show that F has the Darboux property on every interval but that F is not continuous on every interval. Show, too, that F assumes every value in the interval $[-1, 1]$ infinitely often.

Answer ☐

Exercise 107 (fixed points) A function $f : [a, b] \rightarrow [a, b]$ is said to have a fixed point $c \in [a, b]$ if $f(c) = c$. Show that every uniformly continuous function f mapping $[a, b]$ into itself has at least one fixed point.

Answer ☐

Exercise 108 (fixed points) Let $f : [a, b] \rightarrow [a, b]$ be continuous. Define a sequence recursively by $z_1 = x_1$, $z_2 = f(z_1)$, $\dots$, $z_n = f(z_{n-1})$ where $x_1 \in [a, b]$. Show that if the sequence $\{z_n\}$ is convergent, then it must converge to a fixed point of f . Answer $\square$

Exercise 109 Is there a continuous function $f : I \rightarrow \mathbb{R}$ defined on an interval I such that for every real y there are precisely either zero or two solutions to the equation $f(x) = y$? Answer $\square$

Exercise 110 Is there a continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that for every real y there are precisely either zero or three solutions to the equation $f(x) = y$? Answer $\square$

Exercise 111 Suppose that the function $f : \mathbb{R} \rightarrow \mathbb{R}$ is monotone nondecreasing and has the Darboux property. Show that f must be continuous at every point. Answer $\square$

1.8 Derivatives

A derivative² of a function is another function “derived” from the first function by a procedure (which we do not have to review here). Thus, for example, we remember that, if

$$F(x) = x^2 + x + 1$$

then the derived function is

$$F'(x) = 2x + 1.$$

The values of the derived function, $2x + 1$, represent (geometrically) the slope of the tangent line at the points $(x, x^2 + x + 1)$ that are on the graph of the function F . There are numerous other interpretations (other than the geometric) for the values of the derivative function.

Recall the usual notations for derivatives:

$$\frac{d}{dx} \sin x = \cos x.$$

$$F(x) = \sin x, F'(x) = \cos x.$$

²The word derivative in mathematics almost always refers to this concept. In finance, you might have noticed, *derivatives* are financial instrument whose values are “derived” from some underlying security. Observe that the use of the word “derived” is the same.

$$y = \sin x, \quad \frac{dy}{dx} = \cos x.$$

The connection between a function and its derivative is straightforward: the values of the function $F(x)$ are used, along with a limiting process, to determine the values of the derivative function $F'(x)$. That's the definition. We need to know the definition to understand what the derivative signifies, but we do not revert to the definition for computations except very rarely.

The following facts should be familiar:

- A function may or may not have a derivative at a point.
- In order for a function f to have a derivative at a point x_0 the function must be defined at least in some open interval that contains that point.
- A function that has a derivative at a point x_0 is said to be *differentiable* at x_0 . If it fails to have a derivative there then it is said to be *nondifferentiable* at that point.
- There are many calculus tables that can be consulted for derivatives of functions for which familiar formulas are given.
- There are many rules for computation of derivatives for functions that do not appear in the tables explicitly, but for which the tables are nonetheless useful after some further manipulation.
- Information about the derivative function offers deep insight into the nature of the function itself. For example a zero derivative means the function is constant; a nonnegative derivative means the function is increasing. A change in the derivative from positive to negative indicates that a local maximum point in the function was reached.

Exercise 112 ($\epsilon, \delta(x)$ version of derivative) Suppose that F is a differentiable function on an interval I . Show that for every $x \in I$ and every $\epsilon > 0$ there is a $\delta(x) > 0$ so that

$$|F(y) - F(x) - F'(x)(y - x)| \leq \epsilon |y - x|$$

whenever y and x are points in I for which $|y - x| < \delta(x)$.

Answer $\square$

Exercise 113 (differentiable implies continuous) Prove that a function that has a derivative at a point x_0 must also be continuous at that point. Answer $\square$

Exercise 114 ($\epsilon, \delta(x)$ straddled version of derivative) Suppose that F is a differentiable function on an interval I . Show that for every $x \in I$ and every $\epsilon > 0$ there is a $\delta(x) > 0$ so that

$$|F(z) - F(y) - F'(x)(z - y)| \leq \epsilon |z - y|$$

whenever y and z are points in I for which $|y - z| < \delta(x)$ and either $y \leq x \leq z$ or $z \leq x \leq y$. Answer $\square$

Exercise 115 ($\epsilon, \delta(x)$ unstraddled version of derivative) Suppose that F is a differentiable function on an open interval I . Suppose that for every $x \in I$ and every $\epsilon > 0$ there is a $\delta(x) > 0$ so that

$$|F(z) - F(y) - F'(x)(z - y)| \leq \epsilon |z - y|$$

whenever y and z are points in I for which $|y - z| < \delta(x)$ [and we do not require either $y \leq x \leq z$ or $z \leq x \leq y$]. Show that not all differentiable functions would have this property but that if F' is continuous then this property does hold. Answer $\square$

Exercise 116 (locally strictly increasing functions) Suppose that F is a function on an open interval I . Then F is said to be locally strictly increasing at a point x_0 in the interval if there is a $\delta > 0$ so that

$$F(y) < F(x_0) < F(z)$$

for all

$$x_0 - \delta < y < x_0 < z < x_0 + \delta.$$

Show that, if $F'(x_0) > 0$, then F must be locally strictly increasing at x_0 . Show that the converse does not quite hold: if F is differentiable at a point x_0 in the interval and is also locally strictly increasing at x_0 , then necessarily $F'(x_0) \geq 0$ but that $F'(x_0) = 0$ is possible. Answer $\square$

Exercise 117 Suppose that a function F is locally strictly increasing at every point of an open interval (a, b) . Use the Cousin partitioning argument to show that F is strictly increasing on (a, b) .

[In particular, notice that this means that a function with a positive derivative is increasing. This is usually proved using the mean-value theorem that is stated in Section 1.10 below.] Answer $\square$

1.9 Differentiation rules

We remind the reader of the usual calculus formulas by presenting the following slogans. Of course each should be given a precise statement and the proper assumptions clearly made.

Constant rule: if $f(x)$ is constant, then $f' = 0$.

Linear combination rule: $(rf + sg)' = rf' + sg'$ for functions f and g and all real numbers r and s .

Product rule: $(fg)' = f'g + fg'$ for functions f and g .

Quotient rule:

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$$

for functions f and g at points where g does not vanish.

Chain rule: If $f(x) = h(g(x))$, then

$$f'(x) = h'(g(x)) \cdot g'(x).$$

1.10 Mean-value theorem

There is a close connection between the values of a function and the values of its derivative. In one direction this is trivial since the derivative is defined in terms of the values of the function. The other direction is more subtle. How does information about the derivative provide us with information about the function? One of the keys to providing that information is the mean-value theorem.

The usual proof presented in calculus texts requires proving a weak version of the mean-value theorem first (Rolle's theorem) and then using that to prove the full version.

1.10. MEAN-VALUE THEOREM

1.10.1 Rolle's theorem

Theorem 1.22 (Rolle's Theorem) Let f be uniformly continuous on $[a, b]$ and differentiable on (a, b) . If $f(a) = f(b)$ then there must exist at least one point ξ in (a, b) such that $f'(\xi) = 0$.

Exercise 118 Prove the theorem.

Answer ☐

Exercise 119 Interpret the theorem geometrically.

Answer ☐

Exercise 120 Can we claim that the point ξ whose existence is claimed by the theorem, is unique?. How many points might there be?

Answer ☐

Exercise 121 Define a function $f(x) = x \sin x^{-1}$, $f(0) = 0$, on the whole real line. Can Rolle's theorem be applied on the interval $[0, 1/\pi]$?

Answer ☐

Exercise 122 Is it possible to apply Rolle's theorem to the function $f(x) = \sqrt{1-x^2}$ on $[-1, 1]$.

Answer ☐

Exercise 123 Is it possible to apply Rolle's theorem to the function $f(x) = \sqrt{|x|}$ on $[-1, 1]$.

Answer ☐

Exercise 124 Use Rolle's theorem to explain why the cubic equation

$$x^3 + \alpha x^2 + \beta = 0$$

cannot have more than one solution whenever $\alpha > 0$.

Answer ☐

Exercise 125 If the n th-degree equation

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = 0$$

has n distinct real roots, then how many distinct real roots does the $(n-1)$ st degree equation $p'(x) = 0$ have?

Answer ☐

Exercise 126 Suppose that $f'(x) > c > 0$ for all $x \in [0, \infty)$. Show that $\lim_{x \rightarrow \infty} f(x) = \infty$.

Answer ☐

Exercise 127 Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ and both f' and f'' exist everywhere. Show that if f has three zeros, then there must be some point ξ so that $f''(\xi) = 0$. Answer $\square$

Exercise 128 Let f be continuous on an interval $[a, b]$ and differentiable on (a, b) with a derivative that never is zero. Show that f maps $[a, b]$ one-to-one onto some other interval. Answer $\square$

Exercise 129 Let f be continuous on an interval $[a, b]$ and twice differentiable on (a, b) with a second derivative that never is zero. Show that f maps $[a, b]$ two-one onto some other interval; that is, there are at most two points in $[a, b]$ mapping into any one value in the range of f . Answer $\square$

1.10.2 Mean-Value theorem

If we drop the requirement in Rolle's theorem that $f(a) = f(b)$, we now obtain the result that there is a point $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Geometrically, this states that there exists a point $c \in (a, b)$ for which the tangent to the graph of the function at $(c, f(c))$ is parallel to the chord determined by the points $(a, f(a))$ and $(b, f(b))$. (See Figure 1.2.)

This is the mean-value theorem, also known as the law of the mean or the first mean-value theorem (because there are other mean-value theorems).

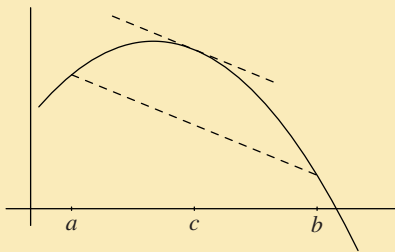
Theorem 1.23 (Mean-Value Theorem) Suppose that f is a continuous function on the closed interval $[a, b]$ and differentiable on (a, b) . Then there exists a point $\xi \in (a, b)$ such that

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Exercise 130 Prove the theorem. Answer $\square$

Exercise 131 Suppose f satisfies the hypotheses of the mean-value theorem on $[a, b]$. Let S be the set of all slopes of chords determined by pairs of points on the graph of f and let

$$D = \{f'(x) : x \in (a, b)\}.$$

Figure 1.2: Mean value theorem [$f'(c)$ is slope of the chord].

1. Prove that $S \subset D$.
2. Give an example to show that D can contain numbers not in S .

Answer ☐

Exercise 132 Interpreting the slope of a chord as an average rate of change and the derivative as an instantaneous rate of change, what does the mean-value theorem say? If a car travels 100 miles in 2 hours, and the position $s(t)$ of the car at time t , measured in hours satisfies the hypotheses of the mean-value theorem, can we be sure that there is at least one instant at which the velocity is 50 mph?

Answer ☐

Exercise 133 Give an example to show that the conclusion of the mean-value theorem can fail if we drop the requirement that f be differentiable at every point in (a,b) .

Answer ☐

Exercise 134 Give an example to show that the conclusion of the mean-value theorem can fail if we drop the requirement of continuity at the endpoints of the interval.

Answer ☐

Exercise 135 Suppose that f is differentiable on $[0, \infty)$ and that

$$\lim_{x \rightarrow \infty} f'(x) = C.$$

Determine

$$\lim_{x \rightarrow \infty} [f(x+a) - f(x)].$$

Answer ☐

Exercise 136 Suppose that f is continuous on $[a, b]$ and differentiable on (a, b) . If

$$\lim_{x \rightarrow a+} f'(x) = C$$

what can you conclude about the right-hand derivative of f at a ?

Answer ☐

Exercise 137 Suppose that f is continuous and that

$$\lim_{x \rightarrow x_0} f'(x)$$

exists. What can you conclude about the differentiability of f ? What can you conclude about the continuity of f' ?

Answer ☐

Exercise 138 Let $f : [0, \infty) \rightarrow \mathbb{R}$ so that f' is decreasing and positive. Show that the series

$$\sum_{i=1}^{\infty} f'(i)$$

is convergent if and only if f is bounded.

Answer ☐

Exercise 139 Prove this second-order version of the mean-value theorem.

Theorem 1.24 (Second order mean-value theorem) Let f be continuous on $[a, b]$ and twice differentiable on (a, b) . Then there exists $c \in (a, b)$ such that

$$f(b) = f(a) + (b-a)f'(a) + (b-a)^2 \frac{f''(c)}{2!}.$$

Answer

☐

Exercise 140 Determine all functions $f : \mathbb{R} \rightarrow \mathbb{R}$ that have the property that

$$f'\left(\frac{x+y}{2}\right) = \frac{f(x) - f(y)}{x-y}$$

1.10. MEAN-VALUE THEOREM

35

for every $x \neq y$.

Answer ☐

Exercise 141 A function is said to be smooth at a point x if

$$\lim_{h \rightarrow 0} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2} = 0.$$

Show that a smooth function need not be continuous. Show that if f'' is continuous at x , then f is smooth at x .

Answer ☐

Exercise 142 Prove this version of the mean-value theorem due to Cauchy.

Theorem 1.25 (Cauchy mean-value theorem) Let f and g be uniformly continuous on $[a, b]$ and differentiable on (a, b) . Then there exists $\xi \in (a, b)$ such that

$$[f(b) - f(a)]g'(\xi) = [g(b) - g(a)]f'(\xi). \quad (1.1)$$

Answer ☐

Exercise 143 Interpret the Cauchy mean-value theorem geometrically.

Answer ☐

Exercise 144 Use Cauchy's mean-value theorem to prove any simple version of L'Hôpital's rule that you can remember from calculus.

Answer ☐

Exercise 145 Show that the conclusion of Cauchy's mean-value can be put into determinant form as

$$\begin{vmatrix} f(a) & g(a) & 1 \\ f(b) & g(b) & 1 \\ f'(c) & g'(c) & 0 \end{vmatrix} = 0.$$

Answer ☐

Exercise 146 Formulate and prove a generalized version of Cauchy's mean-value whose conclusion is the existence of a point c such that

$$\begin{vmatrix} f(a) & g(a) & h(a) \\ f(b) & g(b) & h(b) \\ f'(c) & g'(c) & h'(c) \end{vmatrix} = 0.$$

Answer $\square$

Exercise 147 Suppose that $f : [a, c] \rightarrow \mathbb{R}$ is uniformly continuous and that it has a derivative $f'(x)$ that is monotone increasing on the interval (a, c) . Show that

$$(b - a)f(c) + (c - b)f(a) \geq (c - a)f(b)$$

for any $a < b < c$.

Answer $\square$

Exercise 148 (avoiding the mean-value theorem) The primary use [but not the only use] of the mean-value theorem in a calculus class is to establish that a function with a positive derivative on an open interval (a, b) would have to be increasing. Prove this directly without the easy mean-value proof.

Answer $\square$

Exercise 149 Prove the “converse” to the mean-value theorem:

Let $F, f : [a, b] \rightarrow \mathbb{R}$ and suppose that f is continuous there. Suppose that for every pair of points $a < x < y < b$ there is a point $x < \xi < y$ so that

$$\frac{F(y) - F(x)}{y - x} = f(\xi).$$

Then F is differentiable on (a, b) and f is its derivative.

Answer $\square$

Exercise 150 Let $f : [a, b] \rightarrow \mathbb{R}$ be a uniformly continuous function that is differentiable at all points of the interval (a, b) with possibly finitely many exceptions. Show that there is a point $a < \xi < b$ so that

$$\left| \frac{f(b) - f(a)}{b - a} \right| \leq |f'(\xi)|.$$

Answer $\square$

Exercise 151 (Flett's theorem) Given a function differentiable at every point of an interval $[a, b]$ and with $f'(a) = f'(b)$, show that there is a point ξ in the interval for which

$$\frac{f(\xi) - f(a)}{\xi - a} = f'(\xi).$$

Answer $\square$

1.10.3 The Darboux property of the derivative

We have proved that all continuous functions have the Darboux property. We now prove that all derivatives have the Darboux property. This was proved by Darboux in 1875; one of the conclusions he intended was that there must be an abundance of functions that have the Darboux property and are yet not continuous, since all derivatives have this property and not all derivatives are continuous.

Theorem 1.26 (Darboux property of the derivative) Let F be differentiable on an open interval I . Suppose $a, b \in I$, $a < b$, and $F'(a) \neq F'(b)$. Let γ be any number between $F'(a)$ and $F'(b)$. Then there must exist a point $\xi \in (a, b)$ such that $F'(\xi) = \gamma$.

Exercise 152 Compare Rolle's theorem to Darboux's theorem. Suppose G is everywhere differentiable, that $a < b$ and $G(a) = G(b)$. Then Rolle's theorem asserts the existence of a point ξ in the open interval (a, b) for which $G'(\xi) = 0$. Give a proof of the same thing if the hypothesis $G(a) = G(b)$ is replaced by $G'(a) < 0 < G'(b)$ or $G'(b) < 0 < G'(a)$. Use that to prove Theorem 1.26.

Answer $\square$

Exercise 153 Let $F : \mathbb{R} \rightarrow \mathbb{R}$ be a differentiable function. Show that F' is continuous if and only if the set

$$E_\alpha = \{x : F'(x) = \alpha\}$$

is closed for each real number α .

Answer $\square$

Exercise 154 A function defined on an interval is piecewise monotone if the interval can be subdivided into a finite number of subintervals on each of which the function is nondecreasing or nonincreasing. Show that every polynomial is piecewise monotone.

Answer $\square$

1.10.4 Vanishing derivatives and constant functions

When the derivative is zero we sometimes use colorful language by saying that the derivative *vanishes*! When the derivative of a function vanishes we expect the function to be constant. But how is that really proved?

Theorem 1.27 (vanishing derivatives) *Let $F : [a, b] \rightarrow \mathbb{R}$ be uniformly continuous on the closed, bounded interval $[a, b]$ and suppose that $F'(x) = 0$ for every $a < x < b$. Then F is a constant function on $[a, b]$.*

Corollary 1.28 *Let $F : (a, b) \rightarrow \mathbb{R}$ and suppose that $F'(x) = 0$ for every $a < x < b$. Then F is a constant function on (a, b) .*

Exercise 155 *Prove the theorem using the mean-value theorem.*

Answer ☐

Exercise 156 *Prove the theorem without using the mean-value theorem.*

Answer ☐

Exercise 157 *Deduce the corollary from the theorem.*

Answer ☐

1.10.5 Vanishing derivatives with exceptional sets

When a function has a vanishing derivative then that function must be constant. What if there is a small set of points at which we are unable to determine that the derivative is zero?

Theorem 1.29 (vanishing derivatives with a few exceptions) *Let $F : [a, b] \rightarrow \mathbb{R}$ be uniformly continuous on the closed, bounded interval $[a, b]$ and suppose that $F'(x) = 0$ for every $a < x < b$ with finitely many possible exceptions. Then F is a constant function on $[a, b]$.*

Corollary 1.30 *Let $F : (a, b) \rightarrow \mathbb{R}$ be continuous on the open interval (a, b) and suppose that $F'(x) = 0$ for every $a < x < b$ with finitely many possible exceptions. Then F is a constant function on (a, b) .*

1.10. MEAN-VALUE THEOREM

39

Exercise 158 Prove the theorem by subdividing the interval at the exceptional points.

Answer ☐

Exercise 159 Prove the theorem by applying Exercise 150.

☐

Exercise 160 Prove the corollary.

Answer ☐

Exercise 161 Let $F, G : [a, b] \rightarrow \mathbb{R}$ be uniformly continuous functions on the closed, bounded interval $[a, b]$ and suppose that $F'(x) = f(x)$ for every $a < x < b$ with finitely many possible exceptions, and that $G'(x) = f(x)$ for every $a < x < b$ with finitely many possible exceptions. Show that F and G differ by a constant $[a, b]$.

Answer ☐

Exercise 162 Construct a non-constant function which has a zero derivative at all but finitely many points.

Answer ☐

Exercise 163 Prove the following major improvement of Theorem 1.29. Here, by many exceptions, we include the possibility of infinitely many exceptions provided, only, that it is possible to arrange the exceptional points into a sequence.

Theorem 1.31 (vanishing derivatives with many exceptions) Let $F : [a, b] \rightarrow \mathbb{R}$ be uniformly continuous on the closed, bounded interval $[a, b]$ and suppose that $F'(x) = 0$ for every $a < x < b$ with the possible exception of the points $c_1, c_2, c_3, \dots$ forming an infinite sequence. Show that F is a constant function on $[a, b]$.

[The argument that was successful for Theorem 1.29 will not work for infinitely many exceptional points. A Cousin partitioning argument does work.]

Answer ☐

Exercise 164 Suppose that F is a function continuous at every point of the real line and such that $F'(x) = 0$ for every x that is irrational. Show that F is constant.

Answer ☐

Exercise 165 Suppose that G is a function continuous at every point of the real line and such that $G'(x) = x$ for every x that is irrational. What functions G have such a property?

Answer ☐

Exercise 166 Let $F, G : [a, b] \rightarrow \mathbb{R}$ be uniformly continuous functions on the closed, bounded interval $[a, b]$ and suppose that $F'(x) = f(x)$ for every $a < x < b$ with the possible exception of points in a sequence $\{c_1, c_2, c_3, \dots\}$, and that $G'(x) = f(x)$ for every $a < x < b$ with the possible exception of points in a sequence $\{d_1, d_2, d_3, \dots\}$. Show that F and G differ by a constant $[a, b]$.

Answer ☐

1.11 Lipschitz functions

A function satisfies a Lipschitz condition if there is some limitation on the possible slopes of secant lines, lines joining points $(x, f(x))$ and $(y, f(y))$. Since the slope of such a line would be

$$\frac{f(y) - f(x)}{y - x}$$

any bounds put on this fraction is called a Lipschitz condition.

Definition 1.32 A function f is said to satisfy a Lipschitz condition on an interval I if

$$|f(x) - f(y)| \leq M|x - y|$$

for all x, y in the interval.

Functions that satisfy such a condition are called *Lipschitz functions* and play a key role in many parts of analysis.

Exercise 167 Show that a function that satisfies a Lipschitz condition on an interval must be uniformly continuous on that interval. □

Exercise 168 Show that if f is assumed to be continuous on $[a, b]$ and differentiable on (a, b) then f is a Lipschitz function if and only if the derivative f' is bounded on (a, b) . Answer □

Exercise 169 Show that the function $f(x) = \sqrt{x}$ is uniformly continuous on the interval $[0, \infty)$ but that it does not satisfy a Lipschitz condition on that interval. Answer □

Exercise 170 A function F on an interval I is said to have bounded derived numbers if there is a number M so that, for each $x \in I$ one can choose $\delta > 0$ so that

$$\left| \frac{F(x+h) - F(x)}{h} \right| \leq M$$

whenever $x+h \in I$ and $|h| < \delta$. Using a Cousin partitioning argument, show that F is Lipschitz if and only if F has bounded derived numbers. Answer □

1.11. LIPSCHITZ FUNCTIONS

41

Exercise 171 Is a linear combination of Lipschitz functions also Lipschitz?

Answer ☐

Exercise 172 Is a product of Lipschitz functions also Lipschitz?

Answer ☐

Exercise 173 Is $f(x) = \log x$ a Lipschitz function?

Answer ☐

Exercise 174 Is $f(x) = |x|$ a Lipschitz function?

Answer ☐

Exercise 175 If $F : [a, b] \rightarrow \mathbb{R}$ is a Lipschitz function show that the function $G(x) = F(x) + kx$ is increasing for some value k and decreasing for some other value of k . Is the converse true? ☐

Exercise 176 Show that every polynomial is a Lipschitz function on any bounded interval. What about unbounded intervals? ☐

Exercise 177 In an idle moment a careless student proposed to study a kind of super Lipschitz condition: he supposed that

$$|f(x) - f(y)| \leq M|x - y|^2$$

for all x, y in an interval. What functions would have this property?

Answer ☐

Exercise 178 A function f is said to be bi-Lipschitz on an interval I if there is an $M > 0$ so that

$$\frac{1}{M}|x - y| \leq |f(x) - f(y)| \leq M|x - y|$$

for all x, y in the interval. What can you say about such functions? Can you give examples of such functions? ☐

Exercise 179 Is there a difference between the following two statements:

$$|f(x) - f(y)| < |x - y| \text{ for all } x, y \text{ in an interval}$$

and

$$|f(x) - f(y)| \leq K|x - y| \text{ for all } x, y \text{ in an interval, for some } K < 1?$$

Answer ☐

Exercise 180 If $F_n : [a, b] \rightarrow \mathbb{R}$ is a Lipschitz function for each $n = 1, 2, 3, \dots$ and $F(x) = \lim_{n \rightarrow \infty} F_n(x)$ for each $a \leq x \leq b$, does it follow that F must also be a Lipschitz function. ☐

Chapter 2

The Indefinite Integral

You will, no doubt, remember the formula

$$\int x^2 dx = \frac{x^3}{3} + C$$

from your first calculus classes. This assertion includes the following observations.

- $\frac{d}{dx} \left[\frac{x^3}{3} + C \right] = x^2$.
- Any other function F for which the identity $F'(x) = x^2$ holds is of the form $F(x) = x^3/3 + C$ for some constant C .
- C is called the *constant of integration* and is intended as a completely arbitrary constant.
- The expression $\int x^2 dx$ is intended to be ambiguous and is to include any and all functions whose derivative is x^2 .

In this chapter we will make this rather more precise and we will generalize by allowing a finite exceptional set where the derivative need not exist.

2.1 An indefinite integral on an interval

We will state our definition for open intervals only. We shall assume that indefinite integrals are continuous and we require them to be differentiable everywhere except possibly at a finite set.

Definition 2.1 Let (a, b) be an open interval (bounded or unbounded) and let f be a function defined on that interval except possibly at finitely many points. Then any continuous function $F : (a, b) \rightarrow \mathbb{R}$ for which $F'(x) = f(x)$ for all $a < x < b$ except possibly at finitely many points is said to be an indefinite integral for f on (a, b) .

The notation

$$\int F'(x) dx = F(x) + C$$

will frequently be used. Our indefinite integration theory is essentially the study of continuous functions $F : (a, b) \rightarrow \mathbb{R}$ defined on an open interval, for which there is only a finite number of points of nondifferentiability. Note that, if there are no exceptional points, then we do not have to check that the function is continuous: every differentiable function is continuous.

The indefinite integration theory is, consequently, all about derivatives. We shall see too, in the next chapter, that the *definite* integration theory is also all about derivatives.

Exercise 181 Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is differentiable at every point of the open interval (a, b) . Is F an indefinite integral for F' ? Answer $\square$

Exercise 182 If F is an indefinite integral for a function f on an open interval (a, b) and $a < x < b$, is it necessarily true that $F'(x) = f(x)$. Answer $\square$

Exercise 183 Let $F, G : (a, b) \rightarrow \mathbb{R}$ be two continuous functions for which $F'(x) = f(x)$ for all $a < x < b$ except possibly at finitely many points and $G'(x) = f(x)$ for all $a < x < b$ except possibly at finitely many points. Then F and G must differ by a constant. In particular, on the interval (a, b) the statements

$$\int f(x) dx = F(x) + C_1$$

and

$$\int f(x) dx = G(x) + C_2$$

are both valid (where C_1 and C_2 represent arbitrary constants of integration).

Answer $\square$

2.1.1 Role of the finite exceptional set

The simplest kind of antiderivative is expressed in the situation $F'(x) = f(x)$ for *all* $a < x < b$ [no exceptions]. Our theory is slightly more general in that we allow a finite set of failures and compensate for this by insisting that the function F is continuous at those points.

There is a language that is often adopted to allow exceptions in mathematical statements. We do not use this language in Chapter 2 or Chapter 3 but, for classroom presentation, it might be useful. We will use this language in Chapter 4 and in Part Two of the text.

mostly everywhere A statement holds *mostly everywhere* if it holds everywhere with the exception of a finite set of points $c_1, c_2, c_3, \dots, c_n$.

nearly everywhere A statement holds *nearly everywhere* if it holds everywhere with the exception of a sequence of points $c_1, c_2, c_3, \dots$.

almost everywhere A statement holds *almost everywhere* if it holds everywhere with the exception of a set of measure zero¹.

Thus our indefinite integral is the study of continuous functions that are differentiable mostly everywhere. It is only a *little* bit more ambitious to allow a sequence of points of nondifferentiability. This is the point of view taken in the elementary analysis text, Elias Zakon, *Mathematical Analysis I*, ISBN 1-931705-02-X, published by The Trillia Group, 2004. Thus, in his text, all integrals concern continuous functions that are differentiable nearly everywhere. The mostly everywhere case is the easiest since it needs an appeal only to the mean-value theorem for justification. The nearly everywhere case is rather harder, but if you have worked through the proof of Theorem 1.31 you have seen all the difficulties handled fairly easily.

¹This notion of a set of measure zero will be defined in Chapter 4. For now understand that a set of measure zero is small in a certain sense of measurement.

The final step in the program of improving integration theory is to allow sets of measure zero and study functions that are differentiable almost everywhere. This presents new technical challenges and we shall not attempt it until Chapter 4. Our goal is to get there using Chapters 2 and 3 as elementary warmups.

2.1.2 Features of the indefinite integral

We shall often in the sequel distinguish among the following four cases for an indefinite integral.

Theorem 2.2 *Let F be an indefinite integral for a function f on an open interval (a, b) .*

1. *F is continuous on (a, b) but may or not be uniformly continuous there.*
2. *If f is bounded then F is Lipschitz on (a, b) and hence uniformly continuous there.*
3. *If f unbounded then F is not Lipschitz on (a, b) and may or not be uniformly continuous there.*
4. *If f is nonnegative and unbounded then F is uniformly continuous on (a, b) if and only if F is bounded.*

Exercise 184 *Give an example of two functions f and g possessing indefinite integrals on the interval $(0, 1)$ so that, of the two indefinite integrals F and G , one is uniformly continuous and the other is not.* Answer $\square$

Exercise 185 *Prove this part of Theorem 2.2: If a function f is bounded and possesses an indefinite integral F on (a, b) then F is Lipschitz on (a, b) . Deduce that F is uniformly continuous on (a, b) .* Answer $\square$

2.1.3 The notation $\int f(x) dx$

Since we cannot avoid its use in elementary calculus classes, we define the symbol

$$\int f(x) dx$$

2.1. AN INDEFINITE INTEGRAL ON AN INTERVAL

47

to mean the collection of *all* possible functions that are indefinite integrals of f on an appropriately specified interval. Because of Exercise 183 we know that we can always write this as

$$\int f(x) dx = F(x) + C$$

where F is any one indefinite integral and C is an arbitrary constant called the constant of integration. In more advanced mathematical discussions this notation seldom appears, although there are frequent discussions of indefinite integrals (meaning a function whose derivative is the function being integrated).

Exercise 186 Why exactly is this statement incorrect:

$$\int x^2 dx = x^3/3 + 1?$$

Answer ☐

Exercise 187 Check the identities

$$\frac{d}{dx}(x+1)^2 = 2(x+1)$$

and

$$\frac{d}{dx}(x^2 + 2x) = 2x + 2 = 2(x+1).$$

Thus, on $(-\infty, \infty)$,

$$\int (2x+2) dx = (x+1)^2 + C$$

and

$$\int (2x+2) dx = (x^2 + 2x) + C.$$

Does it follow that $(x+1)^2 = (x^2 + 2x)$?

Answer ☐

Exercise 188 Suppose that we drop continuity from the requirement of an indefinite integral and allow only one point at which the derivative may fail (instead of a finite set of points). Illustrate the situation by finding all possible indefinite integrals [in this new sense] of $f(x) = x^2$ on $(0, 1)$.

Answer ☐

Exercise 189 Show that the function $f(x) = 1/x$ has an indefinite integral on any open interval that does not include zero and does not have an indefinite integral on any open interval containing zero. Is the difficulty here because $f(0)$ is undefined?

Answer ☐

Exercise 190 Show that the function

$$f(x) = \frac{1}{\sqrt{|x|}}$$

has an indefinite integral on any open interval, even if that interval does include zero. Is there any difficulty that arises here because $f(0)$ is undefined?

Answer ☐

Exercise 191 Which is correct

$$\int \frac{1}{x} dx = \log x + C \text{ or } \int \frac{1}{x} dx = \log(-x) + C \text{ or } \int \frac{1}{x} dx = \log |x| + C?$$

Answer ☐

2.2 Existence of indefinite integrals

We cannot be sure in advance that any particular function f has an indefinite integral on a given interval, unless we happen to find one. We turn now to the problem of finding sufficient conditions under which we can be assured that one exists. This is a rather subtle point. Many beginning students might feel that we are seeking to ensure ourselves that an indefinite integral *can be found*. We are, instead, seeking for assurances that an indefinite integral *does indeed exist*. We might still remain completely unable to write down some formula for that indefinite integral because there is no “formula” possible.

We shall show now that, with appropriate continuity assumptions on f , we can be assured that an indefinite integral exists without any requirement that we should find it. Our methods will show that we can also describe a procedure that would, in theory, produce the indefinite integral as the limit of a sequence of simpler functions. This procedure would work only for functions that are mostly continuous. We will still have a theory for indefinite integrals of discontinuous functions but we will have to be content with the fact that much of the theory is formal, and describes objects which are

not necessarily constructible².

2.2.1 Upper functions

We will illustrate our method by introducing the notion of an upper function. This is a piecewise linear function whose slopes dominate the function.

Let f be defined and bounded on an open interval (a, b) and let us choose points

$$a = x_0 < x_1 < x_2 < x_3 < \cdots < x_{n-1} < x_n = b.$$

Suppose that F is a uniformly continuous function on $[a, b]$ that is linear on each interval $[x_{i-1}, x_i]$ and such that

$$\frac{F(x_i) - F(x_{i-1})}{x_i - x_{i-1}} \geq f(\xi)$$

for all $x_{i-1} \leq \xi \leq x_i$ ($i = 1, 2, \dots, n$). Then we can call F an *upper function* for f on $[a, b]$.

The method of upper functions is to approximate the indefinite integral that we require by suitable upper functions. Upper functions are piecewise linear functions with the break points (where the corners are) at $x_1, x_2, \dots, x_{n-1}$. The slopes of these line segments exceed the values of the function f in the corresponding intervals. See Figure 2.1 for an illustration of such a function.

Exercise 192 Let $f(x) = x^2$ be defined on the interval $[0, 1]$. Define an upper function for f using the points

$$0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1.$$

Answer $\square$

Exercise 193 (step functions) Let a function f be defined by requiring that, for any integer n (positive, negative, or zero), $f(x) = n$ if $n - 1 < x < n$. (Values at the integers can be omitted or assigned arbitrary values.) This is a simple example of a step function. Find a formula for an indefinite integral and show that this is an upper function for f .

Answer $\square$

²**Note to the instructor:** Just how unconstructible are indefinite integrals in general? See Chris Freiling, *How to compute antiderivatives*, Bull. Symbolic Logic 1 (1995), no. 3, 279–316. This is by no means an elementary question.

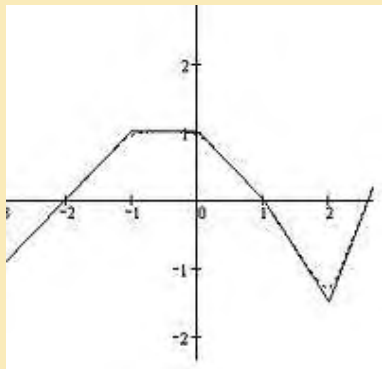


Figure 2.1: A piecewise linear function.

2.2.2 The main existence theorem

For bounded, continuous functions we can always determine an indefinite integral by a limiting process using appropriate upper functions. The theorem is a technical computation that justifies this statement.

Theorem 2.3 Suppose that $f : (a, b) \rightarrow \mathbb{R}$ is a bounded function on an open interval (a, b) [bounded or unbounded]. Then there exists a Lipschitz function $F : (a, b) \rightarrow \mathbb{R}$ so that $F'(x) = f(x)$ for every point $a < x < b$ at which f is continuous.

Corollary 2.4 Suppose that $f : (a, b) \rightarrow \mathbb{R}$ is a bounded function on an open interval (a, b) [bounded or unbounded] and that there are only a finite number of discontinuity points of f in (a, b) . Then f has an indefinite integral on (a, b) , which must be Lipschitz on (a, b) .

Our theorem applies only to bounded functions, but we remember that if f is continuous on (a, b) then it is uniformly continuous, and hence bounded, on any subinterval $[c, d] \subset (a, b)$. This allows the following corollary. Note that we will

not necessarily get an indefinite integral that is Lipschitz on all of (a, b) .

Corollary 2.5 Suppose that $f : (a, b) \rightarrow \mathbb{R}$ is a function on an open interval (a, b) [bounded or unbounded] and that there are no discontinuity points of f in (a, b) . Then f has an indefinite integral on (a, b) .

Exercise 194 Use the method of upper functions to prove Theorem 2.3. It will be enough to assume that $f : (0, 1) \rightarrow \mathbb{R}$ and that f is nonnegative and bounded. (Exercises 195 and 196 ask for the justifications for this assumption.)

Answer ☐

Exercise 195 Suppose that $f : (a, b) \rightarrow \mathbb{R}$ and set $g(t) = f(a + t(b - a))$ for all $0 \leq t \leq 1$. If G is an indefinite integral for g on $(0, 1)$ show how to find an indefinite integral for f on (a, b) .

Answer ☐

Exercise 196 Suppose that $f : (a, b) \rightarrow \mathbb{R}$ is a bounded function and that

$$K = \inf\{f(x) : a < x < b\}.$$

Set $g(t) = f(t) - K$ for all $a < t < b$. Show that g is nonnegative and bounded. Suppose that G is an indefinite integral for g on (a, b) ; show how to find an indefinite integral for f on (a, b) .

Answer ☐

Exercise 197 Show how to deduce Corollary 2.5 from the theorem.

Answer ☐

2.3 Basic properties of indefinite integrals

We conclude our chapter on the indefinite integral by discussing some typical calculus topics. We have developed a precise theory of indefinite integrals and we are beginning to understand the nature of the concept.

There are a number of techniques that have traditionally been taught in calculus courses for the purpose of evaluating or manipulating integrals. Many courses you will take (e.g., physics, applied mathematics, differential equations) will assume that you have mastered these techniques and have little difficulty in applying them.

The reason that you are asked to study these techniques is that they are required for working with integrals or developing theory, not merely for computations. If a course in calculus seems to be overly devoted to evaluating indefinite integrals it is only that you are being drilled in the methods. The skill in finding an exact expression for an indefinite integral is of little use: it won't help in all cases anyway. Besides, any integral that can be handled by these methods can be handled in seconds in Maple or Mathematica (see Section 2.3.5).

2.3.1 Linear combinations

There is a familiar formula for the derivative of a linear combination:

$$\frac{d}{dx} \{rF(x) + sG(x)\} = rF'(x) + sG'(x).$$

This immediately provides a corresponding formula for the indefinite integral of a linear combination:

$$\int (rf(x) + sg(x)) dx = r \int f(x) dx + s \int g(x) dx$$

As usual with statements about indefinite integrals this is only accurate if some mention of an open interval is made. To interpret this formula correctly, let us make it very precise. We assume that both f and g have indefinite integrals F and G on the same interval I . Then the formula claims, merely, that the function $H(x) = rF(x) + sG(x)$ is an indefinite integral of the function $h(x) = rf(x) + sg(x)$ on that interval I .

Exercise 198 (linear combinations) *Prove this formula by showing that $H(x) = rF(x) + sG(x)$ is an indefinite integral of the function $h(x) = rf(x) + sg(x)$ on any interval I , assuming that both f and g have indefinite integrals F and G on I .* Answer $\square$

2.3.2 Integration by parts

There is a familiar formula for the derivative of a product:

$$\frac{d}{dx} \{F(x)G(x)\} = F'(x)G(x) + F(x)G'(x).$$

This immediately provides a corresponding formula for the indefinite integral of a product:

$$\int F(x)G'(x) dx = F(x)G(x) - \int F'(x)G(x) dx.$$

Again we remember that statements about indefinite integrals are only accurate if some mention of an open interval is made. To interpret this formula correctly, let us make it very precise. We assume that $F'G$ has an indefinite integral H on an open interval I . Then the formula claims, merely, that the function $K(x) = F(x)G(x) - H(x)$ is an indefinite integral of the function $F(x)G'(x)$ on that interval I .

Exercise 199 (integration by parts) *Explain and verify the formula.* Answer $\square$

Exercise 200 (calculus student notation) If $u = f(x)$, $v = g(x)$, and we denote $du = f'(x)dx$ and $dv = g'(x)dx$ then in its simplest form the product rule is often described as

$$\int u dv = uv - \int v du.$$

Explain how this version is used.

Answer $\square$

Exercise 201 (extra practise) If you need extra practise on integration by parts as a calculus technique here is a standard collection of examples all cooked in advance so that an integration by parts technique will successfully determine an exact formula for the integral. This is not the case except for very selected examples.

[The interval on which the integration is performed is not specified but it should be obvious which points, if any, to avoid.]

$$\begin{aligned} & \int x e^x dx, \int x \sin x dx, \int x \ln x dx, \int x \cos 3x dx, \int \frac{\ln x}{x^5} dx, \int \arcsin 3x dx, \\ & \int \ln x dx, \int 2x \arctan x dx, \int x^2 e^{3x} dx, \int x^3 \ln 5x dx, \int (\ln x)^2 dx, \int x \sqrt{x+3} dx, \\ & \int x \sin x \cos x dx, \int \left(\frac{\ln x}{x}\right)^2 dx, \int x^5 e^{x^3} dx, \int x^3 \cos(x^2) dx, \int x^7 \sqrt{5+3x^4} dx, \\ & \int \frac{x^3}{(x^2+5)^2} dx, \int e^{6x} \sin(e^{3x}) dx, \int \frac{x^3 e^{x^2}}{(x^2+1)^2} dx, \int e^x \cos x dx \text{ and } \int \sin 3x \cos 5x dx. \end{aligned}$$

Answer $\square$

2.3.3 Change of variable

The chain rule for the derivative of a composition of functions is the formula:

$$\frac{d}{dx} F(G(x)) = F'(G(x)) G'(x).$$

This immediately provides a corresponding formula for the indefinite integral of a product:

$$\int F'(G(x)) G'(x) dx = \int F'(u) du = F(u) + C = F(G(x)) + C \quad [u = G(x)]$$

where we have used the familiar device $u = G(x)$, $du = G'(x)dx$ to make the formula more transparent.

This is called the *change of variable rule*, although it is usually called *integration by substitution* in most calculus presentations.

Again we remember that statements about indefinite integrals are only accurate if some mention of an open interval is made. To interpret this formula correctly, let us make it very precise. We assume that F is a differentiable function on an open interval I . We assume too that G' has an indefinite integral G on an interval J and assumes all of its values in the interval I . Then the formula claims, merely, that the function $F(G(x))$ is an indefinite integral of the function $F'(G(x))G'(x)$ on that interval J [not on the interval I please].

Exercise 202 In the argument for the change of variable rule we did not address the possibility that F might have finitely many points of nondifferentiability. Discuss. Answer $\square$

Exercise 203 Verify that this argument is correct:

$$\int x \cos(x^2 + 1) dx = \frac{1}{2} \int 2x \cos(x^2 + 1) dx = \frac{1}{2} \int \cos u du = \frac{1}{2} \sin u + C = \frac{1}{2} \sin(x^2 + 1) + C.$$

Answer

 $\square$

Exercise 204 Here is a completely typical calculus exercise (or exam question). You are asked to determine an explicit formula for $\int x e^{x^2} dx$. What is expected and how do you proceed? Answer $\square$

Exercise 205 Given that $\int f(t) dt = F(t) + C$ determine $\int f(rx + s) dx$ for any real numbers r and s . Answer $\square$

2.3.4 What is the derivative of the indefinite integral?

What is

$$\frac{d}{dx} \int f(x) dx?$$

By definition this indefinite integral is the family of all functions whose derivative is $f(x)$ [on some pre-specified open interval] but with a possibly finite set of exceptions. So the answer trivially is that

$$\frac{d}{dx} \int f(x) dx = f(x)$$

at most points inside the interval of integration. (But not necessarily at all points.)

The following theorem will do in many situations, but it does not fully answer our question. There are exact derivatives that have very large sets of points at which they are discontinuous.

Theorem 2.6 Suppose that $f : (a, b) \rightarrow \mathbb{R}$ has an indefinite integral F on the interval (a, b) . Then $F'(x) = f(x)$ at every point in (a, b) at which f is continuous.

Exercise 206 Prove the theorem.

Answer $\square$

2.3.5 Partial fractions

Many calculus texts will teach, as an integration tool, the method of partial fractions. It is, actually, an important algebraic technique with applicability in numerous situations, not merely in certain integration problems. It is best to learn this in detail outside of a calculus presentation since it invariably consumes a great deal of student time as the algebraic techniques are tedious at best and, often, reveal a weakness in the background preparation of many of the students.

It will suffice for us to recount the method that will permit the explicit integration of

$$\int \frac{x+3}{x^2-3x-40} dx.$$

The following passage is a direct quotation from the Wikipedia site entry for partial fractions.

“Suppose it is desired to decompose the rational function

$$\frac{x+3}{x^2-3x-40}$$

into partial fractions. The denominator factors as

$$(x-8)(x+5)$$

and so we seek scalars A and B such that

$$\frac{x+3}{x^2-3x-40} = \frac{x+3}{(x-8)(x+5)} = \frac{A}{x-8} + \frac{B}{x+5}.$$

One way of finding A and B begins by "clearing fractions", i.e., multiplying both sides by the common denominator

$(x-8)(x+5)$. This yields

$$x+3 = A(x+5) + B(x-8).$$

Collecting like terms gives

$$x+3 = (A+B)x + (5A-8B).$$

Equating coefficients of like terms then yields:

$$\begin{array}{rclcl} A & + & B & = & 1 \\ 5A & - & 8B & = & 3 \end{array}$$

The solution is $A = 11/13$, $B = 2/13$. Thus we have the partial fraction decomposition

$$\frac{x+3}{x^2-3x-40} = \frac{11/13}{x-8} + \frac{2/13}{x+5} = \frac{11}{13(x-8)} + \frac{2}{13(x+5)}.$$

Alternatively, take the original equation

$$\frac{x+3}{(x-8)(x+5)} = \frac{A}{x-8} + \frac{B}{x+5}.$$

multiply by $(x-8)$ to get

$$\frac{x+3}{x+5} = A + \frac{B(x-8)}{x+5}.$$

Evaluate at $x = 8$ to solve for A as

$$\frac{11}{13} = A.$$

Multiply the original equation by $(x+5)$ to get

$$\frac{x+3}{x-8} = \frac{A(x+5)}{x-8} + B.$$

Evaluate at $x = -5$ to solve for B as

$$\frac{-2}{-13} = \frac{2}{13} = B.$$

As a result of this algebraic identity we can quickly determine that

$$\int \frac{x+3}{x^2-3x-40} dx = [11/13] \log(x-8) + [2/13] \log(x+5) + C.$$

This example is typical and entirely representative of the easier examples that would be expected in a calculus course. The method is, however, much more extensive than this simple computation would suggest. But it is not part of integration theory even if your instructor chooses to drill on it.

Partial fraction method in Maple

Computer algebra packages can easily perform indefinite integration using the partial fraction method without a need for the student to master all the details. Here is a short Maple session illustrating that all the partial fraction details given above are handled easily without resorting to hand calculation. That is not to say that the student should entirely avoid the method itself since it has many theoretical applications beyond its use here.

```
[32]dogwood% maple
|^^|      Maple 12 (SUN SPARC SOLARIS)
_|\\|    |/_|. Copyright (c) Maplesoft, a division of Waterloo Maple Inc. 2008
\ MAPLE / All rights reserved. Maple is a trademark of
<___> Waterloo Maple Inc.
```

```
> int( (x+3) / ( x^2-3*x-40), x);

11
2/13 ln(x + 5) + -- ln(x - 8)
13
```

No constant of integration appears in the result for indefinite integrals.

Exercise 207 *In determining that*

$$\int \frac{x+3}{x^2-3x-40} dx = [11/13] \log(x-8) + [2/13] \log(x+5) + C$$

we did not mention an open interval in which this would be valid. Discuss.

Answer ☐

2.3.6 Tables of integrals

Prior to the availability of computer software packages like Maple³, serious users of the calculus often required access to tables of integrals. If an indefinite integral did have an expression in terms of some formula then it could be found in the tables [if they were extensive enough] or else some transformations using our techniques above (integration by parts, change of variable, etc.) could be applied to find an equivalent integral that did appear in the tables.

Most calculus books (not this one) still have small tables of integrals. Much more efficient, nowadays, is simply to rely on a computer application such as Maple or Mathematica to search for an explicit formula for an indefinite integral. These packages will even tell you if no explicit formula exists.

It is probably a waste of lecture time to teach for long any method that uses tables and it is a waste of paper to write about them. The interested reader should just Google “tables of integrals” to see what can be done. It has the same historical interest that logarithms as devices for computation have. Store your old tables of integrals in the same drawer with your grandparent’s slide rules.

³See especially Section 3.10.1.

Chapter 3

The Definite Integral

We have defined already the notion of an indefinite integral

$$\int F'(x) dx = F(x) + C.$$

The definite integral

$$\int_a^b F'(x) dx = F(b) - F(a)$$

is defined as a special case of that and the connection between the two concepts is immediate. In other calculus courses one might be introduced to a different (also very limited) version of the integral introduced in the middle of the 19th century by Riemann. Then the connection with the indefinite integral is established by means of a deep theorem known as the fundamental theorem of the calculus. Here we run this program backwards. We take the simpler approach of starting with the fundamental theorem as a definition and then recover the Riemann integration methods later.

There are numerous advantages in this. We can immediately start doing some very interesting integration theory and computing integrals. Since we have already learned indefinite integration we have an immediate grasp of the new theory. We are not confined to the limited Riemann integral and we have no need to introduce the improper integral. We can make, eventually, a seamless transition to the Lebesgue integral and beyond.

This calculus integral (also known as “Newton’s integral”) is a limited version of the full integration theory on the real line. It is intended as a teaching method for introducing integration theory. Later, in Chapter 4, we will present an

introduction to the full modern version of integration theory on the real line.

3.1 Definition of the calculus integral

The definite integral is defined directly by means of the indefinite integral and uses a similar notation.

Definition 3.1 *Let f be a function defined at every point of a closed, bounded interval $[a, b]$ with possibly finitely many exceptions. Then f is said to be integrable [calculus sense] if there exists a uniformly continuous function $F : [a, b] \rightarrow \mathbb{R}$ that is an indefinite integral for f on the open interval (a, b) . In that case the number*

$$\int_a^b f(x) dx = F(b) - F(a),$$

is called the definite integral of f on $[a, b]$.

To make this perfectly clear let us recall what this statement would mean. We require:

1. f is defined on $[a, b]$ except possibly at points of a finite set. [In particular $f(a)$ and $f(b)$ need not be defined.]
2. There is a uniformly continuous function F on $[a, b]$,
3. $F'(x) = f(x)$ at every point $a < x < b$ except possibly at points of a finite set.
4. We compute $F(b) - F(a)$ and call this number the definite integral of f on $[a, b]$.

Thus our integration is essentially the study of uniformly continuous functions $F : [a, b] \rightarrow \mathbb{R}$ for which there is only a finite number of points of nondifferentiability. For these functions we use the notation

$$\int_a^b F'(x) dx = F(b) - F(a). \tag{3.1}$$

The integration theory is, consequently, all about derivatives, just as was the indefinite integration theory. The statement (3.1) is here a definition not, as it would be in many other textbooks, a theorem.

3.1.1 Alternative definition of the integral

In many applications it is more convenient to work with a definition that expresses everything within the corresponding open interval (a, b) .

Definition 3.2 Let f be a function defined at every point of a bounded, open interval (a, b) with possibly finitely many exceptions. Then f is said to be integrable [calculus sense] on the closed interval $[a, b]$ if there exists a uniformly continuous indefinite integral F for f on (a, b) . In that case the number

$$\int_a^b f(x) dx = F(b-) - F(a+),$$

is called the definite integral of f on $[a, b]$.

This statement would mean.

1. f is defined at least on (a, b) except possibly at points of a finite set.
2. There is a uniformly continuous function F on (a, b) , with $F'(x) = f(x)$ at every point $a < x < b$ except possibly at points of a finite set.
3. Because F is uniformly continuous on (a, b) , the two one-sided limits

$$\lim_{x \rightarrow a+} F(x) = F(a+) \quad \text{and} \quad \lim_{x \rightarrow b-} F(x) = F(b-)$$

will exist.

4. The number $F(b-) - F(a+)$ is the definite integral of f on $[a, b]$.

Exercise 208 To be sure that a function f is integrable on a closed, bounded interval $[a, b]$ you need to find an indefinite integral F on (a, b) and then check one of the following:

1. F is uniformly continuous on (a, b) , or
2. F is uniformly continuous on $[a, b]$, or

3. F is continuous on (a, b) and the one-sided limits,

$$\lim_{x \rightarrow a+} F(x) = F(a+) \quad \text{and} \quad \lim_{x \rightarrow b-} F(x) = F(b-)$$

exist.

Show that these are equivalent.

Answer $\square$

3.1.2 Infinite integrals

Exactly the same definition for the infinite integrals

$$\int_{-\infty}^{\infty} f(x) dx, \quad \int_a^{\infty} f(x) dx, \quad \text{and} \quad \int_{-\infty}^b f(x) dx$$

can be given as for the integral over a closed bounded interval.

Definition 3.3 Let f be a function defined at every point of (∞, ∞) with possibly finitely many exceptions. Then f is said to be integrable in the calculus sense on (∞, ∞) if there exists an indefinite integral $F : (\infty, \infty) \rightarrow \mathbb{R}$ for f for which both limits

$$F(\infty) = \lim_{x \rightarrow \infty} F(x) \quad \text{and} \quad F(-\infty) = \lim_{x \rightarrow -\infty} F(x)$$

exist. In that case the number

$$\int_{-\infty}^{\infty} f(x) dx = F(\infty) - F(-\infty),$$

is called the definite integral of f on (∞, ∞) .

This statement would mean.

1. f is defined at all real numbers except possibly at points of a finite set.
2. There is a continuous function F on (∞, ∞) , with $F'(x) = f(x)$ at every point except possibly at points of a finite set.

3. The two infinite limits

$$F(\infty) = \lim_{x \rightarrow \infty} F(x) \text{ and } F(-\infty) = \lim_{x \rightarrow -\infty} F(x)$$

exist. This must be checked. For this either compute the limits or else use Exercise 63: for all $\varepsilon > 0$ there should exist a positive number T so that

$$\omega F((T, \infty)) < \varepsilon \text{ and } \omega F((-\infty, T)) < \varepsilon.$$

4. The number $F(\infty) - F(-\infty)$ is the definite integral of f on $[a, b]$.

Similar assertions define

$$\int_{-\infty}^b f(x) dx = F(b) - F(-\infty)$$

and

$$\int_a^{\infty} f(x) dx = F(\infty) - F(a).$$

In analogy with the terminology of an infinite series $\sum_{k=1}^{\infty} a_k$ we often say that the integral $\int_a^{\infty} f(x) dx$ *converges* when the integral exists. That suggests language asserting that the integral *converges absolutely* if both integrals

$$\int_a^{\infty} f(x) dx \text{ and } \int_a^{\infty} |f(x)| dx$$

exist.

3.1.3 Simple properties of integrals

As preliminaries let us state three theorems for this integral that can be proved very quickly just by translating into statements about derivatives. We state these three simple theorems for integrals on bounded intervals but the same methods handle infinite integrals.

Theorem 3.4 (integrability on subintervals) *If f is integrable on a closed, bounded interval $[a, b]$ then f is integrable on any subinterval $[c, d] \subset [a, b]$.*

Theorem 3.5 (additivity of the integral) If f is integrable on the closed, bounded intervals $[a, b]$ and $[b, c]$ then f is integrable on the interval $[a, c]$ and, moreover,

$$\int_a^b f(x) dx + \int_b^c f(x) dx = \int_a^c f(x) dx.$$

Theorem 3.6 (integral inequalities) Suppose that the two functions f, g are both integrable on a closed, bounded interval $[a, b]$ and that $f(x) \leq g(x)$ for all $x \in [a, b]$ with possibly finitely many exceptions. Then

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

Exercise 209 Prove Theorem 3.4 both for integrals on $[a, b]$ or $(-\infty, \infty)$.

Answer ☐

Exercise 210 Prove Theorem 3.5 both for integrals on $[a, b]$ or $(-\infty, \infty)$.

Answer ☐

Exercise 211 Prove Theorem 3.6 both for integrals on $[a, b]$ or $(-\infty, \infty)$.

Answer ☐

Exercise 212 Show that the function $f(x) = x^2$ is integrable on $[-1, 2]$ and compute its definite integral there.

Answer ☐

Exercise 213 Show that the function $f(x) = x^{-1}$ is not integrable on $[-1, 0]$, $[0, 1]$, nor on any closed bounded interval that contains the point $x = 0$. Did the fact that $f(0)$ is undefined influence your argument? Is this function integrable on $(\infty, -1]$ or on $[1, \infty)$?

Answer ☐

Exercise 214 Show that the function $f(x) = x^{-1/2}$ is integrable on $[0, 2]$ and compute its definite integral there. Did the fact that $f(0)$ is undefined interfere with your argument? Is this function integrable on $[0, \infty)$?

Answer ☐

Exercise 215 Show that the function $f(x) = 1/\sqrt{|x|}$ is integrable on any interval $[a, b]$ and determine the value of the integral.

Answer ☐

3.1. DEFINITION OF THE CALCULUS INTEGRAL

65

Exercise 216 (why the finite exceptional set?) In the definition of the calculus integral we permit a finite exceptional set. Why not just skip the exceptional set and just split the interval into pieces? Answer $\square$

Exercise 217 (limitations of the calculus integral) Define a function $F : [0, 1] \rightarrow \mathbb{R}$ in such a way that $F(0) = 0$, and for each odd integer $n = 1, 3, 5, \dots$, $F(1/n) = 1/n$ and each even integer $n = 2, 4, 6, \dots$, $F(1/n) = 0$. On the intervals $[1/(n+1), 1/n]$ for $n = 1, 2, 3$, the function is linear. Show that $\int_a^b F'(x) dx$ exists as a calculus integral for all $0 < a < b \leq 1$ but that $\int_0^1 F'(x) dx$ does not.

Hint: too many exceptional points.

Answer $\square$

Exercise 218 Show that each of the following functions is not integrable on the interval stated:

1. $f(x) = 1$ for all x irrational and $f(x) = 0$ if x is rational, on any interval $[a, b]$.
2. $f(x) = 1$ for all x irrational and $f(x)$ is undefined if x is rational, on any interval $[a, b]$.
3. $f(x) = 1$ for all $x \neq 1, 1/2, 1/3, 1/4, \dots$ and $f(1/n) = c_n$ for some sequence of positive numbers $c_1, c_2, c_3, \dots$, on the interval $[0, 1]$.

Answer $\square$

Exercise 219 Determine all values of p for which the integrals

$$\int_0^1 x^p dx \quad \text{or} \quad \int_1^\infty x^p dx$$

exist.

Answer $\square$

Exercise 220 Are the following additivity formulas for infinite integrals valid:

1. $\int_{-\infty}^\infty f(x) dx = \int_{-\infty}^a f(x) dx + \int_a^b f(x) dx + \int_b^\infty f(x) dx?$
2. $\int_0^\infty f(x) dx = \sum_{n=1}^\infty \int_{n-1}^n f(x) dx?$

$$3. \int_{-\infty}^{\infty} f(x) dx = \sum_{n=-\infty}^{\infty} \int_{n-1}^n f(x) dx?$$

Answer ☐

3.1.4 Integrability of bounded functions

What bounded functions are integrable on an interval $[a, b]$? According to the definition we need to find an indefinite integral on (a, b) and then determine whether it is uniformly continuous. If there is an indefinite integral we know from the mean-value theorem that it would have to be Lipschitz and so must be uniformly continuous. Thus integrability of bounded functions on bounded intervals reduces simply to ensuring that there is an indefinite integral.

Theorem 3.7 *If $f : (a, b) \rightarrow \mathbb{R}$ is a bounded function that is continuous at all but finitely many points of an open bounded interval (a, b) then f is integrable on $[a, b]$.*

Corollary 3.8 *If $f : [a, b] \rightarrow \mathbb{R}$ is a uniformly continuous function then f is integrable on $[a, b]$.*

Exercise 221 *Show that all step functions are integrable.*

Answer ☐

Exercise 222 *Show that all differentiable functions are integrable.*

Answer ☐

Exercise 223 *Show that the Heaviside function is integrable on any interval and show how to compute that integral.* ☐

Exercise 224 *If $f : (a, b) \rightarrow \mathbb{R}$ is a function that is continuous at all points of (a, b) then f is integrable on every closed, bounded subinterval $[c, d] \subset (a, b)$. Show that f is integrable on $[a, b]$ if and only if the one-sided limits*

$$\lim_{t \rightarrow a+} \int_t^c f(x) dx \quad \text{and} \quad \lim_{t \rightarrow b-} \int_c^t f(x) dx$$

exist for any $a < c < b$. Moreover, in that case

$$\int_a^b f(x) dx = \lim_{t \rightarrow a+} \int_t^c f(x) dx + \lim_{t \rightarrow b-} \int_c^t f(x) dx.$$

Answer ☐

3.1.5 Integrability for the unbounded case

What unbounded functions are integrable on an interval $[a, b]$? What functions are integrable on an unbounded interval $(-\infty, \infty)$? Once again we need to find an indefinite integral on (a, b) . The indefinite integral will be continuous but we need to check further details to be sure the definite integral exists. In both cases attention shifts to the endpoints, $F(a+)$ and $F(b-)$ in the case of the integral on $[a, b]$, and $F(-\infty)$ and $F(\infty)$ in the case of the integral on $(-\infty, \infty)$.

The following simple theorems are sometimes called comparison tests for integrals.

Theorem 3.9 (comparison test I) Suppose that $f, g : (a, b) \rightarrow \mathbb{R}$ are functions on (a, b) , both of which have an indefinite integral on (a, b) . Suppose that $|f(x)| \leq g(x)$ for all $a < x < b$. If g is integrable on $[a, b]$ then so too is f .

Theorem 3.10 (comparison test II) Suppose that $f, g : (a, \infty) \rightarrow \mathbb{R}$ are functions on (a, ∞) , both of which have an indefinite integral on (a, ∞) . Suppose that $|f(x)| \leq g(x)$ for all $a < x$. If g is integrable on $[a, \infty)$ then so too is f .

We recall that we know already:

If $f : (a, b) \rightarrow \mathbb{R}$ is an unbounded function that is continuous at all points of (a, b) then f has an indefinite integral on (a, b) . That indefinite integral may or may not be uniformly continuous.

That provides two quick corollaries of our theorems.

Corollary 3.11 Suppose that f is an unbounded function on (a, b) that is continuous at all but a finite number of points, and suppose that $g : (a, b) \rightarrow \mathbb{R}$ with $|f(x)| \leq g(x)$ for all $a < x < b$. If g is integrable on $[a, b]$ then so too is f .

Corollary 3.12 Suppose that f is function on (a, ∞) that is continuous at all but a finite number of points, and suppose that $g : (a, \infty) \rightarrow \mathbb{R}$ with $|f(x)| \leq g(x)$ for all $a < x$. If g is integrable on $[a, \infty)$ then so too is f .

Exercise 225 Prove the two comparison tests [Theorems 3.9 and 3.10].

Answer $\square$

Exercise 226 Prove Corollary 3.11.

Answer ☐

Exercise 227 Prove Corollary 3.12.

Answer ☐

Exercise 228 Which, if any, of these integrals exist:

$$\int_0^{\pi/2} \sqrt{\frac{\sin x}{x}} dx, \quad \int_0^{\pi/2} \sqrt{\frac{\sin x}{x^2}} dx, \quad \text{and} \quad \int_0^{\pi/2} \sqrt{\frac{\sin x}{x^3}} dx?$$

Answer ☐

Exercise 229 Apply the comparison test to each of these integrals:

$$\int_1^{\infty} \frac{\sin x}{\sqrt{x}} dx, \quad \int_1^{\infty} \frac{\sin x}{x} dx, \quad \text{and} \quad \int_1^{\infty} \frac{\sin x}{x^2} dx.$$

Answer ☐

Exercise 230 (nonnegative functions) Show that a nonnegative function $f : (a, b) \rightarrow \mathbb{R}$ is integrable on $[a, b]$ if and only if it has a bounded indefinite integral on (a, b) .

Answer ☐

Exercise 231 Give an example of a function $f : (a, b) \rightarrow \mathbb{R}$ that is not integrable on $[a, b]$ and yet it does have a bounded indefinite integral on (a, b) .

Answer ☐

Exercise 232 Discuss the existence of the definite integral

$$\int_a^b \frac{p(x) dx}{q(x)}$$

where $p(x)$ and $q(x)$ are both polynomials.

Answer ☐

Exercise 233 Discuss the existence of the integral

$$\int_a^{\infty} \frac{p(x)}{q(x)} dx$$

where $p(x)$ and $q(x)$ are polynomials.

Answer ☐

Exercise 234 (the integral test) Let f be a continuous, nonnegative, decreasing function on $[1, \infty)$. Show that the integral $\int_1^\infty f(x) dx$ exists if and only if the series $\sum_{n=1}^\infty f(n)$ converges.

Answer ☐

Exercise 235 Give an example of a function f that is continuous and nonnegative on $[1, \infty)$ so that the integral $\int_1^\infty f(x) dx$ exists but the series $\sum_{n=1}^\infty f(n)$ diverges.

Answer ☐

Exercise 236 Give an example of a function f that is continuous and nonnegative on $[1, \infty)$ so that the integral $\int_1^\infty f(x) dx$ does not exist but the series $\sum_{n=1}^\infty f(n)$ converges.

Answer ☐

3.1.6 Products of integrable functions

When is the product of a pair of integrable functions integrable? When both functions are bounded and defined on a closed, bounded interval we shall likely be successful. When both functions are unbounded, or the interval is unbounded simple examples exist to show that products of integrable functions need not be integrable.

Exercise 237 Suppose we are given a pair of functions f and g such that each is uniformly continuous on $[a, b]$. Show that each of f , g and the product fg is integrable on $[a, b]$.

Answer ☐

Exercise 238 Suppose we are given a pair of functions f and g such that each is bounded and has at most a finite number of discontinuities in (a, b) . Show that each of f , g and the product fg is integrable on $[a, b]$.

Answer ☐

Exercise 239 Find a pair of functions f and g , integrable on $[0, 1]$ and continuous on $(0, 1)$ but such that the product fg is not.

Answer ☐

Exercise 240 Find a pair of continuous functions f and g , integrable on $[1, \infty)$ but such that the product fg is not.

Answer ☐

Exercise 241 Suppose that $F, G : [a, b] \rightarrow \mathbb{R}$ are uniformly continuous functions that are differentiable at all but a finite number of points in (a, b) . Show that $F'G$ is integrable on $[a, b]$ if and only if FG' is integrable on $[a, b]$.

Answer ☐

3.1.7 Notation: $\int_a^a f(x) dx$ and $\int_b^a f(x) dx$

The expressions

$$\int_a^a f(x) dx \text{ and } \int_b^a f(x) dx$$

for $b > a$ do not yet make sense since integration is required to hold on a closed, bounded interval. But these notations are extremely convenient.

Thus we will agree that

$$\int_a^a f(x) dx = 0$$

and, if $a < b$ and the integral $\int_a^b f(x) dx$ exists as a calculus integral, then we assign this meaning to the “backwards” integral:

$$\int_b^a f(x) dx = - \int_a^b f(x) dx.$$

Exercise 242 Suppose that the integral $\int_a^b f(x) dx$ exists as a calculus integral and that F is an indefinite integral for f on that interval. Does the formula

$$\int_s^t f(x) dx = F(t) - F(s) \quad (s, t \in [a, b])$$

work even if $s = t$ or if $s > t$?

Answer ☐

Exercise 243 Check that the formula

$$\int_a^b f(x) dx + \int_b^c f(x) dx = \int_a^c f(x) dx$$

works for all real numbers a , b and c .

Answer ☐

3.1.8 The dummy variable: what is the “ x ” in $\int_a^b f(x) dx$?

If you examine the two statements

$$\int x^2 dx = x^3/3 + C \text{ and } \int_1^2 x^2 dx = 2^3/3 - 1^3/3 = 7/3$$

you might notice an odd feature. The first integral [the indefinite integral] requires the symbol x to express the functions on both sides. But in the second integral [the definite integral] the symbol x plays no role except to signify the function being integrated. If we had given the function a name, say $g(x) = x^2$ then the first identity could be written

$$\int g(x) dx = x^3/3 \quad \text{or} \quad \int g(t) dt = t^3/3 + c$$

while the second one might be more simply written as

$$\int_1^2 g = 7/3.$$

In definite integrals the symbols x and dx are considered as dummy variables, useful for notational purposes and helpful as aids to computation, but carrying no significance. Thus you should feel free [and are encouraged] to use any other letters you like to represent the dummy variable. But do not use a letter that serves some other purpose elsewhere in your discussion.

Here are some bad and even terrible abuses of this:

... let $x = 2$ and let $y = \int_1^2 x^2 dx$.

... Show that $\int_1^x x^2 dx = x^3/3 - 1/3$.

Exercise 244 *Do you know of any other bad uses of dummy variables?*

Answer $\square$

3.1.9 Definite vs. indefinite integrals

The connection between the definite and indefinite integrals is immediate; we have simply defined one in terms of the other.

If F is an indefinite integral of an integrable function f on an interval (a, b) then

$$\int_a^b f(x) dx = F(b-) - F(a+).$$

In the other direction if f is integrable on an interval $[a, b]$ then, on the open interval (a, b) , the indefinite integral can be expressed as

$$\int f(x) dx = \int_a^x f(t) dt + C.$$

Both statements are tautologies; this is a matter of definition not of computation or argument.

Exercise 245 A student is asked to find the indefinite integral of e^{2x} and he writes

$$\int e^{2x} dx = \int_0^x e^{2t} dt + C.$$

How would you grade?

Answer ☐

Exercise 246 A student is asked to find the indefinite integral of e^{x^2} and she writes

$$\int e^{x^2} dx = \int_0^x e^{t^2} dt + C.$$

How would you grade?

Answer ☐

3.1.10 The calculus student's notation

The procedure that we have learned to compute a definite integral is actually just the definition. For example, if we wish to evaluate

$$\int_{-5}^6 x^2 dx$$

we first determine that

$$\int x^2 dx = x^3/3 + C$$

on any interval. So that, using the function $F(x) = x^3/3$ as an indefinite integral,

$$\int_{-5}^6 x^2 dx = F(6) - F(-5) = 6^3/3 - (-5)^3/3 = (6^3 + 5^3)/3.$$

Calculus students often use a shortened notation for this computation:

$$\int_{-5}^6 x^2 dx = \left. \frac{x^3}{3} \right|_{x=-5}^{x=6} = 6^3/3 - (-5)^3/3.$$

Exercise 247 Would you accept this notation:

$$\int_1^\infty \frac{dx}{\sqrt{x^3}} = -\frac{2}{\sqrt{x}} \Big|_{x=1}^{x=\infty} = 0 - (-2)?$$

Answer $\square$

3.2 Mean-value theorems for integrals

In general the expression

$$\frac{1}{b-a} \int_a^b f(x) dx$$

is thought of as an averaging operation on the function f , determining its “average value” throughout the whole interval $[a, b]$. The *first mean-value theorem* for integrals says that the function actually attains this average value at some point inside the interval, i.e., under appropriate hypotheses there is a point $a < \xi < b$ at which

$$\frac{1}{b-a} \int_a^b f(x) dx = f(\xi).$$

But this is nothing new to us. Since the integral is defined by using an indefinite integral F for f this is just the observation that

$$\frac{1}{b-a} \int_a^b f(x) dx = \frac{F(b) - F(a)}{b-a} = f(\xi),$$

the very familiar mean-value theorem for derivatives.

Theorem 3.13 Let $f : (a, b) \rightarrow \mathbb{R}$ be integrable on $[a, b]$ and suppose that F is an indefinite integral. Suppose further that $F'(x) = f(x)$ for all $a < x < b$ with no exceptional points. Then there must exist a point $\xi \in (a, b)$ so that

$$\int_a^b f(x) dx = f(\xi)(b-a).$$

Corollary 3.14 Let $f : (a, b) \rightarrow \mathbb{R}$ be integrable on $[a, b]$ and suppose that f is continuous at each point of (a, b) . Then there must exist a point $\xi \in (a, b)$ so that

$$\int_a^b f(x) dx = f(\xi)(b - a).$$

Exercise 248 Give an example of an integrable function for which the first mean-value theorem for integrals fails.

Answer $\square$

Exercise 249 (another mean-value theorem) Suppose that $G : [a, b] \rightarrow \mathbb{R}$ is a continuous function and $\varphi : [a, b] \rightarrow \mathbb{R}$ is an integrable, nonnegative function. If $G(t)\varphi(t)$ is integrable, show that there exists a number $\xi \in (a, b)$ such that

$$\int_a^b G(t)\varphi(t) dt = G(\xi) \int_a^b \varphi(t) dt.$$

Answer $\square$

Exercise 250 (and another) Suppose that $G : [a, b] \rightarrow \mathbb{R}$ is a positive, monotonically decreasing function and $\varphi : [a, b] \rightarrow \mathbb{R}$ is an integrable function. Suppose that $G\varphi$ is integrable. Then there exists a number $\xi \in (a, b]$ such that

$$\int_a^b G(t)\varphi(t) dt = G(a+0) \int_a^\xi \varphi(t) dt.$$

Note: Here, as usual, $G(a+0)$ stands for $\lim_{x \rightarrow a+} G(x)$, the existence of which follows from the monotonicity of the function G . Note that ξ in the exercise might possibly be b . $\square$

Exercise 251 (... and another) Suppose that $G : [a, b] \rightarrow \mathbb{R}$ is a monotonic (not necessarily decreasing and positive) function and $\varphi : [a, b] \rightarrow \mathbb{R}$ is an integrable function. Suppose that $G\varphi$ is integrable. Then there exists a number $\xi \in (a, b)$ such that

$$\int_a^b G(t)\varphi(t) dt = G(a+0) \int_a^\xi \varphi(t) dt + G(b-0) \int_\xi^b \varphi(t) dt.$$

$\square$

Exercise 252 (Dirichelet integral) As an application of mean-value theorems, show that the integral

$$\int_0^{\infty} \frac{\sin x}{x} dx$$

is convergent but is not absolutely convergent.

Answer $\square$

3.3 Riemann sums

If $F : [a, b] \rightarrow \mathbb{R}$ is a uniformly continuous function that is differentiable at every point of the open interval (a, b) [i.e., every point with no exceptions] then we know that $f = F'$ is integrable and that the first mean-value theorem can be applied to express the integral in the form

$$\int_a^b f(x) dx = F(b) - F(a) = f(\xi)(b - a)$$

for some $\xi \in (a, b)$.

Take any point $a < x_1 < b$ and do the same thing in both of the intervals $[a, x_1]$ and $[x_1, b]$. Then

$$\int_a^b f(x) dx = f(\xi_1)(x_1 - a) + f(\xi_2)(b - x_1)$$

for some points $\xi_1 \in (a, x_1)$ and $\xi_2 \in (x_1, b)$.

In fact then we can do this for any number of points. Let

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

Then there must exist points $\xi_i \in (x_{i-1}, x_i)$ for $i = 1, 2, \dots, n$ so that

$$\int_a^b f(x) dx = \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}). \quad (3.2)$$

We express this observation in the language of partitions and Riemann sums.¹

¹These sums and the connection with integration theory do not originate with Riemann nor are they that late in the history of the subject. Poisson in 1820 proposed such an investigation as “the fundamental proposition of the theory of definite integrals.” Euler by at least 1768 had already used such sums to approximate integrals. Of course, for both of them the integral was understood in our sense as an antiderivative. See

Definition 3.15 (Riemann sum) Suppose that

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, \dots, n\}$$

is a partition of an interval $[a, b]$ and that a function f is defined at every point of the interval $[a, b]$. Then any sum of the form

$$\sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}).$$

is called a Riemann sum for the function f .

Using this language, we have just proved in the identity (3.2) that an integral in many situations can be computed exactly by some Riemann sum. This seems both wonderful and, maybe, not so wonderful. In the first place it means that an integral $\int_a^b f(x) dx$ can be computed by a simple sum using the values of the function f rather than by using the definition and having, instead, to solve a difficult or impossible indefinite integration problem. On the other hand this only works if we can select the right points $\{\xi_i\}$.

3.3.1 Exact computation by Riemann sums

We have proved the following theorem that shows that, in some situations, the definite integral can be computed exactly by a Riemann sum. The proof is obtained directly from the first mean-value theorem for integrals, which itself is simply the mean-value theorem for derivatives.

Judith V. Grabiner, *Who gave you the epsilon? Cauchy and the origins of rigorous calculus*, American Mathematical Monthly 90 (3), 1983, 185–194.

Theorem 3.16 Let $f : (a, b) \rightarrow \mathbb{R}$ be integrable on $[a, b]$ and suppose that F is an indefinite integral. Suppose further that $F'(x) = f(x)$ for all $a < x < b$ with the possible exception of points in a finite set $C \subset (a, b)$. Choose any points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

so that at least all points of C are included. Then there must exist points $\xi_i \in (x_{i-1}, x_i)$ for $i = 1, 2, \dots, n$ so that

$$\int_{x_{i-1}}^{x_i} f(x) dx = f(\xi_i)(x_i - x_{i-1}) \quad (i = 1, 2, 3, \dots, n)$$

and

$$\int_a^b f(x) dx = \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}).$$

Exercise 253 Show that the integral $\int_a^b x dx$ can be computed exactly by any Riemann sum

$$\int_a^b x dx = \sum_{i=1}^n \frac{x_i + x_{i-1}}{2} (x_i - x_{i-1}) = \frac{1}{2} \sum_{i=1}^n (x_i^2 - x_{i-1}^2).$$

Answer ☐

Exercise 254 Subdivide the interval $[0, 1]$ at the points $x_0 = 0$, $x_1 = 1/3$, $x_2 = 2/3$ and $x_3 = 1$. Determine the points ξ_i so that

$$\int_0^1 x^2 dx = \sum_{i=1}^3 \xi_i^2 (x_i - x_{i-1}).$$

☐

Exercise 255 Subdivide the interval $[0, 1]$ at the points $x_0 = 0$, $x_1 = 1/3$, $x_2 = 2/3$ and $x_3 = 1$. Determine the points $\xi_i \in [x_{i-1}, x_i]$ so that

$$\sum_{i=1}^3 \xi_i^2 (x_i - x_{i-1})$$

is as large as possible. By how much does this sum exceed $\int_0^1 x^2 dx$?

☐

Exercise 256 Subdivide the interval $[0, 1]$ at the points $x_0 = 0, x_1 = 1/3, x_2 = 2/3$ and $x_3 = 1$. Consider various choices of the points $\xi_i \in [x_{i-1}, x_i]$ in the sum

$$\sum_{i=1}^3 \xi_i^2 (x_i - x_{i-1}).$$

What are all the possible values of this sum? What is the relation between this set of values and the number $\int_0^1 x^2 dx$? $\square$

Exercise 257 Subdivide the interval $[0, 1]$ by defining the points $x_0 = 0, x_1 = 1/n, x_2 = 2/n, \dots, x_{n-1} = (n-1)/n$, and $x_n = n/n = 1$. Determine the points $\xi_i \in [x_{i-1}, x_i]$ so that

$$\sum_{i=1}^n \xi_i^2 (x_i - x_{i-1}).$$

is as large as possible. By how much does this sum exceed $\int_0^1 x^2 dx$? $\square$

Exercise 258 Let $0 < r < 1$. Subdivide the interval $[0, 1]$ by defining the points $x_0 = 0, x_1 = r^{n-1}, x_2 = r^{n-2}, \dots, x_{n-1} = r^{n-(n-1)} = r$, and $x_n = r^{n-(-n)} = 1$. Determine the points $\xi_i \in [x_{i-1}, x_i]$ so that

$$\sum_{i=1}^n \xi_i^2 (x_i - x_{i-1}).$$

is as large as possible. By how much does this sum exceed $\int_0^1 x^2 dx$? $\square$

Exercise 259 (error estimate) Let $f : [a, b] \rightarrow \mathbb{R}$ be an integrable function. Suppose further that $F'(x) = f(x)$ for all $a < x < b$ where F is an indefinite integral. Suppose that

$$\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\}$$

is an arbitrary partition of $[a, b]$. Show that

$$\left| \int_{x_{i-1}}^{x_i} f(x) dx - f(\xi_i)(x_i - x_{i-1}) \right| \leq \omega f([x_i, x_{i-1}]) (x_i - x_{i-1}) \quad (i = 1, 2, 3, \dots, n)$$

and that

$$\left| \int_a^b f(x) dx - \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \right| \leq \sum_{i=1}^n \omega f([x_i, x_{i-1}]) (x_i - x_{i-1}). \quad (3.3)$$

Note: that, if the right hand side of the inequality (3.3) is small then the Riemann sum, while not precisely equal to the integral, would be a good estimate. Of course, the right hand side might also be big. Answer $\square$

3.3.2 Uniform Approximation by Riemann sums

While Theorem 3.16 shows that all calculus integrals can be exactly computed by Riemann sums, it gives no procedure for determining the correct partition. Suppose we relax our goal. Instead of asking for an exact computation perhaps an approximate computation might be useful:

$$\int_a^b f(x) dx \approx \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1})?$$

By a *uniform* approximation we mean that we shall specify the smallness of the intervals $[x_i, x_{i-1}]$ by a single small number δ . In Section 3.3.4 we specify this smallness in a more general way, by requiring that the length of $[x_i, x_{i-1}]$ be smaller than $\delta(\xi_i)$. This is the *pointwise* version.

Theorem 3.17 *Let f be a bounded function that is defined and continuous at every point of (a, b) with at most finitely many exceptions: Then, f is integrable on $[a, b]$ and moreover the integral may be uniformly approximated by Riemann sums: for every $\varepsilon > 0$ there is a $\delta > 0$ so that*

$$\sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f(x) dx - f(\xi_i)(x_i - x_{i-1}) \right| < \varepsilon$$

and

$$\left| \int_a^b f(x) dx - \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \right| < \varepsilon$$

whenever $\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\}$ is a partition of $[a, b]$ with each

$$x_i - x_{i-1} < \delta$$

and $\xi_i \in [x_{i-1}, x_i]$ is a point where f is defined.

Corollary 3.18 Let $f : [a, b] \rightarrow \mathbb{R}$ be a uniformly continuous function. Then, f is integrable on $[a, b]$ and moreover the integral may be uniformly approximated by Riemann sums.

Exercise 260 Prove Theorem 3.17 in the case when f is uniformly continuous on $[a, b]$ by using the error estimate in Exercise 259. Answer ☐

Exercise 261 Prove Theorem 3.17 in the case when f is continuous on (a, b) . Answer ☐

Exercise 262 Complete the proof of Theorem 3.17. Answer ☐

Exercise 263 Let $f : [a, b] \rightarrow \mathbb{R}$ be an integrable function on $[a, b]$ and suppose, moreover, the integral may be uniformly approximated by Riemann sums. Show that f would have to be bounded. Answer ☐

Exercise 264 Show that the integral

$$\int_0^1 x^2 dx = \lim_{n \rightarrow \infty} \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 + \cdots + n^2}{n^3}.$$

Answer ☐

Exercise 265 Show that the integral

$$\int_0^1 x^2 dx = \lim_{r \rightarrow 1^-} [(1-r) + r(r-r^2) + r^2(r^2-r^3) + r^3(r^3-r^4) + \cdots].$$

Answer ☐

Exercise 266 Show that the integral $\int_0^1 x^5 dx$ can be exactly computed by the method of Riemann sums provided one has the formula

$$1^5 + 2^5 + 3^5 + 4^5 + 5^5 + 6^5 + \cdots + N^5 = \frac{N^6}{6} + \frac{N^5}{2} + \frac{5N^4}{12} - \frac{N^2}{12}.$$

☐

3.3.3 Theorem of G. A. Bliss

Students of the calculus and physics are often required to “set up” integrals by which is meant interpreting a problem as an integral. Basically this amounts to interpreting the problem as a limit of Riemann sums

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}).$$

In this way the student shows that the integral captures all the computations of the problem. In simple cases this is easy enough, but complications can arise.

For example if f and g are two continuous functions, sometimes the correct set up would involve a sum of the form

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i)g(\eta_i)(x_i - x_{i-1})$$

and not the more convenient

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i)g(\xi_i)(x_i - x_{i-1}).$$

Here, rather than a single point ξ_i associated with the interval $[x_i, x_{i-1}]$, two different points ξ_i and η_i must be used.

Nineteenth century students had been taught a rather murky method for handling this case known as the Duhamel principle; it involved an argument using infinitesimals that, at bottom, was simply manipulations of Riemann sums. Bliss² felt that this should be clarified and so produced an elementary theorem of which Theorem 3.19 is a special case. It is just a minor adjustment to our Theorem 3.17.

²G. A. Bliss, *A substitute for Duhamel's theorem*, Annals of Mathematics, Ser. 2, Vol. 16, (1914).

Theorem 3.19 (Bliss) Let f and g be bounded functions that are defined and continuous at every point of (a, b) with at most finitely many exceptions: Then, fg is integrable on $[a, b]$ and moreover the integral may be uniformly approximated by Riemann sums in this alternative sense: for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$\left| \int_a^b f(x)g(x) dx - \sum_{i=1}^n f(\xi_i)g(\xi_i^*)(x_i - x_{i-1}) \right| < \varepsilon$$

whenever $\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\}$ is a partition of $[a, b]$ with each

$$x_i - x_{i-1} < \delta$$

and $\xi_i, \xi_i^* \in [x_{i-1}, x_i]$ with both ξ_i and ξ_i^* points in (a, b) where fg is defined.

Exercise 267 Prove the Bliss theorem.

Answer $\square$

Exercise 268 Prove this further variant of Theorem 3.17.

Theorem 3.20 (Bliss) Let $f_1, f_2, \dots, f_p$ bounded functions that are defined and continuous at every point of (a, b) with at most finitely many exceptions: Then, the product $f_1 f_2 f_3 \dots f_p$ is integrable on $[a, b]$ and moreover the integral may be uniformly approximated by Riemann sums in this alternative sense: for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$\left| \int_a^b f_1(x) f_2(x) f_3(x) \dots f_p(x) dx - \sum_{i=1}^n f_1(\xi_i) f_2(\xi_i^{(2)}) f_3(\xi_i^{(3)}) \dots f_p(\xi_i^{(p)}) (x_i - x_{i-1}) \right| < \varepsilon$$

whenever $\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\}$ is a partition of $[a, b]$ with each

$$x_i - x_{i-1} < \delta$$

and $\xi_i, \xi_i^{(2)}, \xi_i^{(3)}, \dots, \xi_i^{(p)} \in [x_{i-1}, x_i]$ with these being points in (a, b) where the functions are defined.

Answer $\square$ **Exercise 269** Prove one more variant of Theorem 3.17.**Theorem 3.21** Suppose that the function $H(s, t)$ satisfies

$$|H(s, t)| \leq M(|s| + |t|)$$

for some real number M and all real numbers s and t . Let f and g be bounded functions that are defined and continuous at every point of (a, b) with at most finitely many exceptions: Then, $H(f(x), g(x))$ is integrable on $[a, b]$ and moreover the integral may be uniformly approximated by Riemann sums in this sense: for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$\left| \int_a^b H(f(x), g(x)) dx - \sum_{i=1}^n H(f(\xi_i), g(\xi_i^*)) (x_i - x_{i-1}) \right| < \varepsilon$$

whenever $\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\}$ is a partition of $[a, b]$ with each

$$x_i - x_{i-1} < \delta$$

and $\xi_i, \xi_i^* \in [x_{i-1}, x_i]$ with both ξ_i and ξ_i^* points in (a, b) where f and g are defined.

Answer $\square$

3.3.4 Pointwise approximation by Riemann sums

For unbounded, but integrable, functions there cannot be a *uniform* approximation by Riemann sums. Even for bounded functions there will be no uniform approximation by Riemann sums unless the function is mostly continuous; we have proved one direction and later will be able to characterize those functions permitting a uniform approximation as functions that are bounded and “almost everywhere” continuous.

If we are permitted to adjust the smallness of the partition in a pointwise manner, however, then such an approximation by Riemann sums is available. This is less convenient, of course, since for each ε we need find not merely a single positive δ but a positive δ at each point of the interval. While this appears, at the outset, to be a deep property of calculus integrals it is an entirely trivial property.

Much more remarkable is that Henstock,³ who first noted the property, was able also to recognize that all Lebesgue integrable functions have the same property and that this property characterized the much more general integral of Denjoy and Perron. Thus we will see this property again, but next time it will appear as a condition that is both necessary and sufficient.

Theorem 3.22 (Henstock property) *Let $f : [a, b] \rightarrow \mathbb{R}$ be defined and integrable on $[a, b]$. Then, for every $\varepsilon > 0$ and for each point x in $[a, b]$ there is a $\delta(x) > 0$ so that*

$$\sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f(x) dx - f(\xi_i)(x_i - x_{i-1}) \right| < \varepsilon$$

and

$$\left| \int_a^b f(x) dx - \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \right| < \varepsilon$$

whenever $\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\}$ is a partition of $[a, b]$ with each

$$x_i - x_{i-1} < \delta(\xi_i) \quad \text{and} \quad \xi_i \in [x_{i-1}, x_i].$$

Note that our statement requires that the function f being integrated is defined at all points of the interval $[a, b]$. This is not really an inconvenience since we could simply set $f(x) = 0$ (or any other value) at points where the given function f is not defined. The resulting integral is indifferent to changing the value of a function at finitely many points.

Note also that, if there are *no* such partitions having the property of the statements in Theorem 3.22, then the statement is certainly valid, but has no content. This is not the case, i.e., no matter what choice of a function $\delta(x)$ occurs in this situation there must be at least one partition having this property. This is precisely the Cousin covering argument.

Exercise 270 *In the statement of the theorem show that if the first inequality*

$$\sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f(x) dx - f(\xi_i)(x_i - x_{i-1}) \right| < \varepsilon$$

³Ralph Henstock (1923-2007) first worked with this concept in the 1950s while studying nonabsolute integration theory. The characterization of the Denjoy-Perron integral as a pointwise limit of Riemann sums was at the same time discovered by the Czech mathematician Jaroslav Kurzweil and today that integral is called the Henstock-Kurzweil integral by most users.

holds then the second inequality

$$\left| \int_a^b f(x) dx - \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \right| < \varepsilon$$

must follow by simple arithmetic.

Answer ☐

Exercise 271 Prove Theorem 3.22 in the case when f is the exact derivative of an everywhere differentiable function F .

Answer ☐

Exercise 272 Prove Theorem 3.17 in the case where F is an everywhere differentiable function except at one point c inside (a, b) at which F is continuous.

Answer ☐

Exercise 273 Complete the proof of Theorem 3.17.

Answer ☐

3.4 Properties of the integral

The basic properties of integrals are easily obtained for us because the integral is defined directly by differentiation. Thus we can apply all the rules we know about derivatives to obtain corresponding facts about integrals.

3.4.1 Inequalities

Formula for *inequalities*:

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx$$

if $f(x) \leq g(x)$ for all but finitely many points x in (a, b) .

Here is a precise statement of what we intend by this statement: If both functions $f(x)$ and $g(x)$ have a calculus integral on the interval $[a, b]$ and, if $f(x) \leq g(x)$ for all but finitely many points x in (a, b) , then the stated inequality must hold.

The proof is an easy exercise in derivatives. We know that if H is uniformly continuous on $[a, b]$ and if

$$\frac{d}{dx} H(x) \geq 0$$

for all but finitely many points x in (a, b) then $H(x)$ must be nondecreasing on $[a, b]$.

Exercise 274 Complete the details needed to prove the inequality formula.

Answer $\square$

3.4.2 Linear combinations

Formula for *linear combinations*:

$$\int_a^b [rf(x) + sg(x)] dx = r \int_a^b f(x) dx + s \int_a^b g(x) dx \quad (r, s \in \mathbb{R}).$$

Here is a precise statement of what we intend by this formula: If both functions $f(x)$ and $g(x)$ have a calculus integral on the interval $[a, b]$ then any linear combination $rf(x) + sg(x)$ ($r, s \in \mathbb{R}$) also has a calculus integral on the interval $[a, b]$ and, moreover, the identity must hold. The proof is an easy exercise in derivatives. We know that

$$\frac{d}{dx} (rF(x) + sG(x)) = rF'(x) + sG'(x)$$

at any point x at which both F and G are differentiable.

Exercise 275 Complete the details needed to prove the linear combination formula.

$\square$

3.4.3 Subintervals

Formula for *subintervals*: If $a < c < b$ then

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx. \tag{3.4}$$

The intention of the formula is contained in two statements in this case:

If the function $f(x)$ has a calculus integral on the interval $[a, b]$ then $f(x)$ must also have a calculus integral on any closed subinterval of the interval $[a, b]$ and, moreover, the identity (3.4) must hold.

and

If the function $f(x)$ has a calculus integral on the interval $[a, c]$ and also on the interval $[c, b]$ then $f(x)$ must also have a calculus integral on the interval $[a, b]$ and, moreover, the identity (3.4) must hold.

Exercise 276 Supply the details needed to prove the subinterval formula. □

3.4.4 Integration by parts

Integration by parts formula:

$$\int_a^b F(x)G'(x) dx = F(x)G(x) - \int_a^b F'(x)G(x) dx$$

The intention of the formula is contained in the product rule for derivatives:

$$\frac{d}{dx} (F(x)G(x)) = F(x)G'(x) + F'(x)G(x)$$

which holds at any point where both functions are differentiable. One must then give strong enough hypotheses that the function $F(x)G(x)$ is an indefinite integral for the function

$$F(x)G'(x) + F'(x)G(x)$$

in the sense needed for our integral.

Exercise 277 Supply the details needed to prove the integration by parts formula in the special case where F and G are continuously differentiable everywhere. □

Exercise 278 Supply the details needed to state and prove an integration by parts formula that is stronger than the one in the preceding exercise. □

3.4.5 Change of variable

The *change of variable* formula (i.e., integration by substitution):

$$\int_a^b f(g(t))g'(t) dt = \int_{g(a)}^{g(b)} f(x) dx.$$

The intention of the formula is contained in the following statement which contains a sufficient condition that allows this formula to be proved: Let I be an interval and $g : [a, b] \rightarrow I$ a continuously differentiable function. Suppose that

$F : I \rightarrow \mathbb{R}$ is an integrable function. Then the function $F(g(t))g'(t)$ is integrable on $[a, b]$ and the function f is integrable on the interval $[g(a), g(b)]$ (or rather on $[g(a), g(b)]$ if $g(b) < g(a)$) and the identity holds. There are various assumptions under which this might be valid.

The proof is an application of the chain rule for the derivative of a composite function:

$$\frac{d}{dx}F(G(x)) = F'(G(x))G'(x).$$

Exercise 279 Supply the details needed to prove the change of variable formula in the special case where F and G are differentiable everywhere. Answer $\square$

Exercise 280 (a failed change of variables) Let $F(x) = |x|$ and $G(x) = x^2 \sin x^{-1}$, $G(0) = 0$. Does

$$\int_0^1 F'(G(x))G'(x) dx = F(G(1)) - F(G(0)) = |\sin 1|?$$

Answer

 $\square$

Exercise 281 (calculus student notation) Explain the procedure being used by this calculus student:

In the integral $\int_0^2 x \cos(x^2 + 1) dx$ we substitute $u = x^2 + 1$, $du = 2x dx$ and obtain

$$\int_0^2 x \cos(x^2 + 1) dx = \frac{1}{2} \int_{u=1}^{u=5} \cos u du = \frac{1}{2} (\sin(5) - \sin(1)).$$

$\square$

Exercise 282 (calculus student notation) Explain the procedure being used by this calculus student:

The substitution $x = \sin u$, $dx = \cos u du$ is useful, because $\sqrt{1 - \sin^2 u} = \cos u$. Therefore

$$\int_0^1 \sqrt{1 - x^2} dx = \int_0^{\frac{\pi}{2}} \sqrt{1 - \sin^2 u} \cos u du = \int_0^{\frac{\pi}{2}} \cos^2 u du.$$

$\square$

Exercise 283 Supply the details needed to prove the change of variable formula in the special case where G is strictly increasing and differentiable everywhere. Answer $\square$

Exercise 284 Show that the integral

$$\int_0^{\pi^2} \frac{\cos \sqrt{x}}{\sqrt{x}} dx$$

exists and use a change of variable to determine the exact value.

Answer $\square$

3.4.6 What is the derivative of the definite integral?

What is

$$\frac{d}{dx} \int_a^x f(t) dt?$$

We know that $\int_a^x f(t) dt$ is an indefinite integral of f and so, by definition,

$$\frac{d}{dx} \int_a^x f(t) dt = f(x)$$

at all but finitely many points in the interval (a, b) if f is integrable on $[a, b]$.

If we need to know more than that then there is the following version which we have already proved:

$$\frac{d}{dx} \int_a^x f(t) dt = f(x)$$

at all points $a < x < b$ at which f is continuous. We should keep in mind, though, that there may also be many points where f is discontinuous and yet the derivative formula holds.

Advanced note. If we go beyond the calculus interval, as we do in Chapter 4, then the same formula is valid

$$\frac{d}{dx} \int_a^x f(t) dt = f(x)$$

but there may be many more than finitely many exceptions possible. For “most” values of t this is true but there may even be infinitely many exceptions possible. It will still be true at points of continuity but it must also be true at most points when an integrable function is badly discontinuous (as it may well be).

3.5 Absolute integrability

If a function f is integrable, does it necessarily follow that the absolute value of that function, $|f|$, is also integrable? This is important in many applications. Since a solution to this problem rests on the concept of the total variation of a function, we will give that definition below in Section 3.5.1.

Definition 3.23 (absolutely integrable) A function f is absolutely integrable on an interval $[a, b]$ if both f and $|f|$ are integrable there.

Exercise 285 Show that, if f is absolutely integrable on an interval $[a, b]$ then

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

Answer $\square$

Exercise 286 (preview of bounded variation) Show that if a function f is absolutely integrable on a closed, bounded interval $[a, b]$ and F is its indefinite integral then, for all choices of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b,$$

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| \leq \int_a^b |f(x)| dx.$$

Answer $\square$

Exercise 287 (calculus integral is a nonabsolute integral) An integration method is an absolute integration method if whenever a function f is integrable on an interval $[a, b]$ then the absolute value $|f|$ is also integrable there. Show that the calculus integral is a nonabsolute integration method⁴.

Hint: Consider

$$\frac{d}{dx} x \cos\left(\frac{\pi}{x}\right).$$

Answer $\square$

⁴Both the Riemann integral and the Lebesgue integral are absolute integration methods.

3.5. ABSOLUTE INTEGRABILITY

91

Exercise 288 Repeat Exercise 287 but using

$$\frac{d}{dx} x^2 \sin\left(\frac{1}{x^2}\right).$$

Show that this derivative exists at every point. Thus there is an exact derivative which is integrable on every interval but not absolutely integrable. Answer $\square$

Exercise 289 Let f be continuous at every point of (a, b) with at most finitely many exceptions and suppose that f is bounded. Show that f is absolutely integrable on $[a, b]$. Answer $\square$

3.5.1 Functions of bounded variation

The clue to the property that expresses absolute integrability is in Exercise 286. The notion is due to Jordan and the language is that of variation, meaning here a measurement of how much the function is fluctuating.

Definition 3.24 (total variation) A function $F : [a, b] \rightarrow \mathbb{R}$ is said to be of bounded variation if there is a number M so that

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| \leq M$$

for all choices of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

The least such number M is called the total variation of F on $[a, b]$ and is written $V(F, [a, b])$. If F is not of bounded variation then we set $V(F, [a, b]) = \infty$.

Definition 3.25 (total variation function) let $F : [a, b] \rightarrow \mathbb{R}$ be a function of bounded variation. Then the function

$$T(x) = V(F, [a, x]) \quad (a < x \leq b), \quad T(a) = 0$$

is called the total variation function for F on $[a, b]$.

Our main theorem in this section establishes the properties of the total variation function and gives, at least for continuous functions, the connection this concept has with absolute integrability.

Theorem 3.26 (properties of the total variation) *let $F : [a, b] \rightarrow \mathbb{R}$ be a function of bounded variation and let $T(x) = V(F, [a, x])$ be its total variation. Then*

1. *for all $a \leq c < d \leq b$,*

$$|F(d) - F(c)| \leq V(F, [c, d]) = T(d) - T(c).$$

2. *T is monotonic, nondecreasing on $[a, b]$.*

3. *If F is continuous at a point $a < x_0 < b$ then so too is T .*

4. *If F is uniformly continuous on $[a, b]$ then so too is T .*

5. *If F is continuously differentiable at a point $a < x_0 < b$ then so too is T and, moreover $T'(x_0) = |F'(x_0)|$.*

6. *If F is uniformly continuous on $[a, b]$ and continuously differentiable at all but finitely many points in (a, b) then F' is absolutely integrable and*

$$F(x) - F(a) = \int_a^x F'(t) dt \text{ and } T(x) = \int_a^x |F'(t)| dt.$$

As we see here in assertion (6.) of the theorem and will discover further in the exercises, the two notions of total variation and absolute integrability are closely interrelated. The notion of total variation plays such a significant role in the study of real functions in general and in integration theory in particular that it is worthwhile spending some considerable time on it, even at an elementary calculus level. Since the ideas are closely related to other ideas which we are studying this topic should seem a natural development of the theory. Indeed we will find that our discussion of arc length in Section 3.9.3 will require a use of this same language.

Exercise 290 *Show directly from the definition that if $F : [a, b] \rightarrow \mathbb{R}$ is a function of bounded variation then F is a bounded function on $[a, b]$.* Answer $\square$

3.5. ABSOLUTE INTEGRABILITY

93

Exercise 291 Compute the total variation for a function F that is monotonic on $[a, b]$.

Answer ☐

Exercise 292 Compute the total variation function for the function $F(x) = \sin x$ on $[-\pi, \pi]$.

Answer ☐

Exercise 293 Let $F(x)$ have the value zero everywhere except at the point $x = 0$ where $F(0) = 1$. Choose points

$$-1 = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = 1.$$

What are all the possible values of

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})|?$$

What is $V(F, [-1, 1])$?

Answer ☐

Exercise 294 Give an example of a function F defined everywhere and with the property that $V(F, [a, b]) = \infty$ for every interval $[a, b]$.

Answer ☐

Exercise 295 Show that if $F : [a, b] \rightarrow \mathbb{R}$ is Lipschitz then F is a function of bounded variation. Is the converse true?

Answer ☐

Exercise 296 Show that $V(F + G, [a, b]) \leq V(F, [a, b]) + V(G, [a, b])$.

Answer ☐

Exercise 297 Does

$$V(F + G, [a, b]) = V(F, [a, b]) + V(G, [a, b]).$$

Answer ☐

Exercise 298 Prove Theorem 3.26.

Answer ☐

Exercise 299 (Jordan decomposition) Show that a function F has bounded variation on an interval $[a, b]$ if and only if it can be expressed as the difference of two monotonic, nondecreasing functions.

Answer ☐

Exercise 300 Show that the function $F(x) = x \cos\left(\frac{\pi}{x}\right)$, $F(0) = 0$ is continuous everywhere but does not have bounded variation on the interval $[0, 1]$, i.e., that $V(F, [0, 1]) = \infty$.

Answer ☐

Exercise 301 (derivative of the variation) Suppose that $F(x) = x^r \cos x^{-1}$ for $x > 0$, $F(x) = -(-x)^r \cos x^{-1}$ for $x < 0$, and finally $F(0) = 0$. Show that if $r > 1$ then F has bounded variation on $[-1, 1]$ and that $F'(0) = 0$. Let T be the total variation function of F . Show that $T'(0) = 0$ if $r > 2$, that $T'(0) = 2/\pi$ if $r = 2$, and that $T'(0) = \infty$ if $1 < r < 2$.

Note: In particular, at points where F is differentiable, the total variation T need not be. Theorem 3.26 said, in contrast, that at points where F is continuously differentiable, the total variation T must also be continuously differentiable.

Answer □

Exercise 302 (uniformly approximating the variation) Suppose that F is uniformly continuous on $[a, b]$. Show that for any $v < V(F, [a, b])$ there is a $\delta > 0$ so that so that

$$v < \sum_{i=1}^n |F(x_i) - F(x_{i-1})| \leq V(F, [a, b])$$

for all choices of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

provided that each $x_i - x_{i-1} < \delta$. Is it possible to drop or relax the assumption that F is continuous?

Note: This means the variation of a continuous function can be computed much like our Riemann sums approximation to the integral.

Answer □

Exercise 303 Let $F_k : [a, b] \rightarrow \mathbb{R}$ ($k = 1, 2, 3, \dots$) be a sequence of functions of bounded variation, suppose that

$$F(x) = \lim_{k \rightarrow \infty} F_k(x)$$

for every $k = 1, 2, 3, \dots$ and suppose that there is a number M so that

$$V(F_k, [a, b]) \leq M. \quad k = 1, 2, 3, \dots$$

Show that F must also have bounded variation.

Does this prove that every limit of a sequence of functions of bounded variation must also have bounded variation? □

Exercise 304 (locally of bounded variation) Let $F : \mathbb{R} \rightarrow \mathbb{R}$ be a function. We say that F is locally of bounded variation at a point x if there is some positive δ so that $V(F, [x - \delta, x + \delta]) < \infty$. Show that F has bounded variation on every compact interval $[a, b]$ if and only if F is locally of bounded variation at every point $x \in \mathbb{R}$.

Answer □

3.5.2 Indefinite integrals and bounded variation

In the preceding section we spent some time mastering the important concept of total variation. We now see that it precisely describes the absolute integrability of a function. Indefinite integrals of nonabsolutely integrable functions will not be of bounded variation; indefinite integrals of absolutely integrable functions must be of bounded variation.

Theorem 3.27 Suppose that a function $f : (a, b) \rightarrow \mathbb{R}$ is absolutely integrable on a closed, bounded interval $[a, b]$. Then its indefinite integral F must be a function of bounded variation there and, moreover,

$$V(F, [a, b]) = \int_a^b |f(x)| dx.$$

This theorem states only a necessary condition for absolute integrability. If we add in a continuity assumption we can get a complete picture of what happens. Continuity is needed for the calculus integral, but is not needed for more advanced theories of integration.

Theorem 3.28 Let $F : [a, b] \rightarrow \mathbb{R}$ be a uniformly continuous function that is continuously differentiable at every point in a bounded, open interval (a, b) with possibly finitely many exceptions. Then F' is integrable on $[a, b]$ and will be, moreover, absolutely integrable on $[a, b]$ if and only if F has bounded variation on that interval.

Exercise 305 Prove Theorem 3.27.

Answer ☐

Exercise 306 Prove Theorem 3.28.

Answer ☐

3.6 Sequences and series of integrals

Throughout the 18th century much progress in applications of the calculus was made through quite liberal use of the formulas

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \left\{ \lim_{n \rightarrow \infty} f_n(x) \right\} dx$$

and

$$\sum_{k=1}^{\infty} \int_a^b g_k(x) dx = \int_a^b \left\{ \sum_{k=1}^{\infty} g_k(x) \right\} dx.$$

These are vitally important tools but they require careful application and justification. That justification did not come until the middle of the 19th century.

We introduce two definitions of convergence allowing us to interpret what the limit and sum of a sequence,

$$\lim_{n \rightarrow \infty} f_n(x) \quad \text{and} \quad \sum_{k=1}^{\infty} g_k(x)$$

should mean. We will find that uniform convergence allows an easy justification for the basic formulas above. Point-wise convergence is equally important but more delicate. At the level of a calculus course we will find that uniform convergence is the concept we shall use most frequently.

3.6.1 The counterexamples

We begin by asking, naively, whether there is any difficulty in taking limits in the calculus. Suppose that $f_1, f_2, f_3, \dots$ is a sequence of functions defined on an open interval $I = (a, b)$. We suppose that this sequence *converges pointwise* to a function f , i.e., that for each $x \in I$ the sequence of numbers $\{f_n(x)\}$ converges to the value $f(x)$.

Is it true that

1. If f_n is bounded on I for all n , then is f also bounded on I ?
2. If f_n is continuous on I for all n , then is f also continuous on I ?
3. If f_n is uniformly continuous on I for all n , then is f also uniformly continuous on I ?
4. If f_n is differentiable on I for all n , then is f also differentiable on I and, if so, does

$$f' = \lim_{n \rightarrow \infty} f'_n?$$

5. If f_n is integrable on a subinterval $[c, d]$ of I for all n , then is f also integrable on $[c, d]$ and, if so, does

$$\lim_{n \rightarrow \infty} \int_c^d f_n(x) dx = \int_c^d \left\{ \lim_{n \rightarrow \infty} f_n(x) \right\} dx?$$

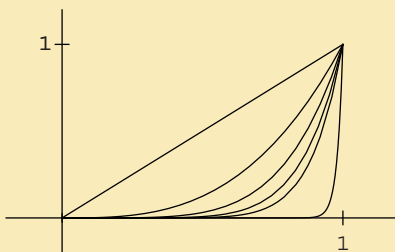


Figure 3.1: Graphs of x^n on $[0, 1]$ for $n = 1, 3, 5, 7, 9$, and 50 .

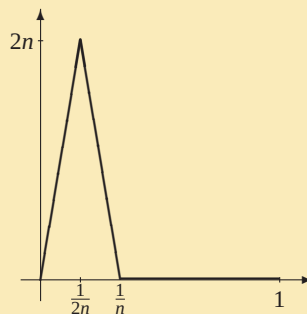
These five questions have *negative* answers in general, as the examples that follow show.

Exercise 307 (An unbounded limit of bounded functions) On the interval $(0, \infty)$ and for each integer n let $f_n(x) = 1/x$ for $x > 1/n$ and $f_n(x) = n$ for each $0 < x \leq 1/n$. Show that each function f_n is both continuous and bounded on $(0, \infty)$. Is the limit function $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ also continuous? Is the limit function bounded? Answer $\square$

Exercise 308 (A discontinuous limit of continuous functions) For each integer n and $-1 < x \leq 1$, let $f_n(x) = x^n$. For $x > 1$ let $f_n(x) = 1$. Show that each f_n is a continuous function on $(-1, \infty)$ and that the sequence converges pointwise to a function f on $(-1, \infty)$ that has a single point of discontinuity. Answer $\square$

Exercise 309 (A limit of uniformly continuous functions) Show that the previous exercise supplies a pointwise convergence sequence of uniformly continuous functions on the interval $[0, 1]$ that does not converge to a uniformly continuous function. $\square$

Exercise 310 (The derivative of the limit is not the limit of the derivative) Let $f_n(x) = x^n/n$ for $-1 < x \leq 1$ and let $f_n(x) = x - (n-1)/n$ for $x > 1$. Show that each f_n is differentiable at every point of the interval $(-1, \infty)$ but that the limit function has a point of nondifferentiability. Answer $\square$

Figure 3.2: Graph of $f_n(x)$ on $[0, 1]$ in Exercise 311.

Exercise 311 (The integral of the limit is not the limit of the integrals) In this example we consider a sequence of continuous functions, each of which has the same integral over the interval. For each n let f_n be defined on $[0, 1]$ as follows: $f_n(0) = 0$, $f_n(1/(2n)) = 2n$, $f_n(1/n) = 0$, f_n is linear on $[0, 1/(2n)]$ and on $[1/(2n), 1/n]$, and $f_n = 0$ on $[1/n, 1]$. (See Figure 3.2.)

It is easy to verify that $f_n \rightarrow 0$ on $[0, 1]$. Now, for each n ,

$$\int_0^1 f_n x = 1.$$

But

$$\int_0^1 (\lim_{n \rightarrow \infty} n f_n(x)) dx = \int_0^1 0 dx = 0.$$

Thus

$$\lim_{n \rightarrow \infty} n \int_0^1 f_n x \neq \int_0^1 \lim_{n \rightarrow \infty} n f_n(x) dx$$

so that the limit of the integrals is not the integral of the limit. $\square$

Exercise 312 (interchange of limit operations) To prove the (false) theorem that the pointwise limit of a sequence of

3.6. SEQUENCES AND SERIES OF INTEGRALS

99

continuous functions is continuous, why cannot we simply write

$$\lim_{x \rightarrow x_0} \left(\lim_{n \rightarrow \infty} f_n(x) \right) = \lim_{n \rightarrow \infty} f_n(x_0) = \lim_{n \rightarrow \infty} \left(\lim_{x \rightarrow x_0} f_n(x) \right)$$

and deduce that

$$\lim_{x \rightarrow x_0} f(x) = f(x_0)?$$

This assumes f_n is continuous at x_0 and “proves” that f is continuous at x_0 . that

Answer $\square$

Exercise 313 Is there anything wrong with this “proof” that a limit of bounded functions is bounded? If each f_n is bounded on an interval I then there must be, by definition, a number M so that $|f_n(x)| \leq M$ for all x in I . By properties of sequence limits

$$|f(x)| = \left| \lim_{n \rightarrow \infty} f_n(x) \right| \leq M$$

also, so f is bounded.

Answer $\square$

Exercise 314 (interchange of limit operations) Let

$$S_{mn} = \begin{cases} 0, & \text{if } m \leq n \\ 1, & \text{if } m > n. \end{cases}$$

Viewed as a matrix,

$$[S_{mn}] = \begin{bmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 1 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

where we are placing the entry S_{mn} in the m th row and n th column. Show that

$$\lim_{n \rightarrow \infty} \left(\lim_{m \rightarrow \infty} S_{mn} \right) \neq \lim_{m \rightarrow \infty} \left(\lim_{n \rightarrow \infty} S_{mn} \right).$$

Answer $\square$

Exercise 315 Examine the pointwise limiting behavior of the sequence of functions

$$f_n(x) = \frac{x^n}{1+x^n}.$$

□

Exercise 316 Show that the natural logarithm function can be expressed as the pointwise limit of a sequence of “simpler” functions,

$$\log x = \lim_{n \rightarrow \infty} n (\sqrt[n]{x} - 1)$$

for every point in the interval. If the answer to our initial five questions for this particular limit is affirmative, what can you deduce about the continuity of the logarithm function? What would be its derivative? What would be $\int_1^2 \log x \, dx$? □

Exercise 317 Let $x_1, x_2, \dots$ be a sequence that contains every rational number, let

$$f_n(x) = \begin{cases} 1, & \text{if } x \in \{x_1, \dots, x_n\} \\ 0, & \text{otherwise,} \end{cases} \quad \text{and} \quad f(x) = \begin{cases} 1, & \text{if } x \text{ is rational} \\ 0, & \text{otherwise.} \end{cases}$$

1. Show that $f_n \rightarrow f$ pointwise on any interval.
2. Show that f_n has only finitely many points of discontinuity while f has no points of continuity.
3. Show that each f_n has a calculus integral on any interval $[c, d]$ while f has a calculus integral on no interval.
4. Show that, for any interval $[c, d]$,

$$\lim_{n \rightarrow \infty} \int_c^d f_n(x) \, dx \neq \int_c^d \left\{ \lim_{n \rightarrow \infty} f_n(x) \right\} \, dx.$$

Answer □

Exercise 318 Let $f_n(x) = \sin nx / \sqrt{n}$. Show that $\lim_{n \rightarrow \infty} n f_n = 0$ but $\lim_{n \rightarrow \infty} n f'_n(0) = \infty$. □

Exercise 319 Let $f_n \rightarrow f$ pointwise at every point in the interval $[a, b]$. We have seen that even if each f_n is continuous it does not follow that f is continuous. Which of the following statements are true?

1. If each f_n is increasing on $[a, b]$, then so is f .
2. If each f_n is nondecreasing on $[a, b]$, then so is f .

3.6. SEQUENCES AND SERIES OF INTEGRALS

101

3. If each f_n is bounded on $[a, b]$, then so is f .
4. If each f_n is everywhere discontinuous on $[a, b]$, then so is f .
5. If each f_n is constant on $[a, b]$, then so is f .
6. If each f_n is positive on $[a, b]$, then so is f .
7. If each f_n is linear on $[a, b]$, then so is f .
8. If each f_n is convex on $[a, b]$, then so is f .

Answer $\square$

Exercise 320 A careless student⁵ once argued as follows: “It seems to me that one can construct a curve without a tangent in a very elementary way. We divide the diagonal of a square into n equal parts and construct on each subdivision as base a right isosceles triangle. In this way we get a kind of delicate little saw. Now I put $n = \infty$. The saw becomes a continuous curve that is infinitesimally different from the diagonal. But it is perfectly clear that its tangent is alternately parallel now to the x -axis, now to the y -axis.” What is the error? (Figure 3.3 illustrates the construction.) Answer $\square$

Exercise 321 Consider again the sequence $\{f_n\}$ of functions $f_n(x) = x^n$ on the interval $(0, 1)$. We saw that $f_n \rightarrow 0$ pointwise on $(0, 1)$, and we proved this by establishing that, for every fixed $x_0 \in (0, 1)$ and $\varepsilon > 0$,

$$|x_0|^n < \varepsilon \text{ if and only if } n > \log \varepsilon / \log x_0.$$

Is it possible to find an integer N so that, for all $x \in (0, 1)$,

$$|x|^n < \varepsilon \text{ if } n > N?$$

Discuss.

Answer $\square$

⁵In this case the “careless student” was the great Russian analyst N. N. Luzin (1883–1950), who recounted in a letter [reproduced in *Amer. Math. Monthly*, 107, (2000), pp. 64–82] how he offered this argument to his professor after a lecture on the Weierstrass continuous nowhere differentiable function.

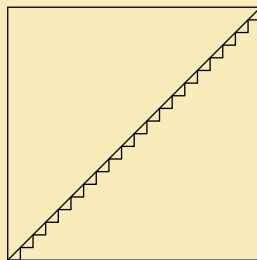


Figure 3.3: Construction in Exercise 320.

3.6.2 Uniform convergence

The most immediate of the conditions which allows an interchange of limits in the calculus is the notion of uniform convergence. This is a very much stronger condition than pointwise convergence.

Definition 3.29 Let $\{f_n\}$ be a sequence of functions defined on an interval I . We say that $\{f_n\}$ converges uniformly to a function f on I if, for every $\varepsilon > 0$, there exists an integer N such that

$$|f_n(x) - f(x)| < \varepsilon \text{ for all } n \geq N \text{ and all } x \in I.$$

Exercise 322 Show that the sequence of functions $f_n(x) = x^n$ converges uniformly on any interval $[0, \eta]$ provided that $0 < \eta < 1$. Answer $\square$

Exercise 323 Using this definition of the Cauchy Criterion

Definition 3.30 (Cauchy Criterion) Let $\{f_n\}$ be a sequence of functions defined on an interval set I . The sequence is said to be uniformly Cauchy on I if for every $\varepsilon > 0$ there exists an integer N such that if $n \geq N$ and $m \geq N$, then $|f_m(x) - f_n(x)| < \varepsilon$ for all $x \in I$.

3.6. SEQUENCES AND SERIES OF INTEGRALS

103

prove the following theorem:

Theorem 3.31 Let $\{f_n\}$ be a sequence of functions defined on an interval I . Then there exists a function f defined on the interval I such that the sequence uniformly on I if and only if $\{f_n\}$ is uniformly Cauchy on I .

Answer $\square$

Exercise 324 In Exercise 322 we showed that the sequence $f_n(x) = x^n$ converges uniformly on any interval $[0, \eta]$, for $0 < \eta < 1$. Prove this again, but using the Cauchy criterion.

Answer $\square$

Exercise 325 (Cauchy criterion for series) The Cauchy criterion can be expressed for uniformly convergent series too. We say that a series $\sum_{k=1}^{\infty} g_k$ converges uniformly to the function f on an interval I if the sequence of partial sums $\{S_n\}$ where

$$S_n(x) = \sum_{k=1}^n g_k(x)$$

converges uniformly to f on I . Prove this theorem:

Theorem 3.32 Let $\{g_k\}$ be a sequence of functions defined on an interval I . Then the series $\sum_{k=1}^{\infty} f_k$ converges uniformly to some function f on the interval I if and only if for every $\varepsilon > 0$ there is an integer N so that

$$\left| \sum_{j=m}^n f_j(x) \right| < \varepsilon$$

for all $n \geq m \geq N$ and all $x \in I$.

Answer $\square$

Exercise 326 Show that the series

$$1 + x + x^2 + x^3 + x^4 + \dots$$

converges pointwise on $[0, 1)$, converges uniformly on any interval $[0, \eta]$ for $0 < \eta < 1$, but that the series does not converge uniformly on $[0, 1)$.

Answer $\square$

Exercise 327 (Weierstrass M-Test) Prove the following theorem, which is usually known as the Weierstrass M-test for uniform convergence of series.

Theorem 3.33 (M-Test) Let $\{f_k\}$ be a sequence of functions defined on an interval I and let $\{M_k\}$ be a sequence of positive constants. If

$$\sum_{k=1}^{\infty} M_k < \infty \quad \text{and} \quad |f_k(x)| \leq M_k \quad \text{for each } x \in I \text{ and } k = 0, 1, 2, \dots,$$

then the series $\sum_{k=1}^{\infty} f_k$ converges uniformly on the interval I .

Answer ☐

Exercise 328 Consider again the geometric series $1 + x + x^2 + \dots$ (as we did in Exercise 326). Use the Weierstrass M-test to prove uniform convergence on the interval $[-a, a]$, for any $0 < a < 1$.

Answer ☐

Exercise 329 Use the Weierstrass M-test to investigate the uniform convergence of the series

$$\sum_{k=1}^{\infty} \frac{\sin k\theta}{k^p}$$

on an interval for values of $p > 0$.

Answer ☐

Exercise 330 (Abel's Test for Uniform Convergence) Prove Abel's test for uniform convergence:

Theorem 3.34 (Abel) Let $\{a_k\}$ and $\{b_k\}$ be sequences of functions on an interval I . Suppose that there is a number M so that

$$-M \leq s_N(x) = \sum_{k=1}^N a_k(x) \leq M$$

for all $x \in I$ and every integer N . Suppose that the sequence of functions $\{b_k\} \rightarrow 0$ converges monotonically to zero at each point and that this convergence is uniform on I . Then the series $\sum_{k=1}^{\infty} a_k(x)b_k(x)$ converges uniformly on I .

Answer ☐

3.6. SEQUENCES AND SERIES OF INTEGRALS

105

Exercise 331 Apply Theorem 3.34, to the following series that often arises in Fourier analysis:

$$\sum_{k=1}^{\infty} \frac{\sin k\theta}{k}.$$

Answer ☐

Exercise 332 Examine the uniform limiting behavior of the sequence of functions

$$f_n(x) = \frac{x^n}{1+x^n}.$$

On what sets can you determine uniform convergence?

☐

Exercise 333 Examine the uniform limiting behavior of the sequence of functions

$$f_n(x) = x^2 e^{-nx}.$$

On what sets can you determine uniform convergence? On what sets can you determine uniform convergence for the sequence of functions $n^2 f_n(x)$?

☐

Exercise 334 Prove that if $\{f_n\}$ and $\{g_n\}$ both converge uniformly on an interval I , then so too does the sequence $\{f_n + g_n\}$.

☐

Exercise 335 Prove or disprove that if $\{f_n\}$ and $\{g_n\}$ both converge uniformly on an interval I , then so too does the sequence $\{f_n g_n\}$.

☐

Exercise 336 Prove or disprove that if f is a continuous function on $(-\infty, \infty)$, then

$$f(x + 1/n) \rightarrow f(x)$$

uniformly on $(-\infty, \infty)$. (What extra condition, stronger than continuity, would work if not?)

☐

Exercise 337 Prove that $f_n \rightarrow f$ converges uniformly on an interval I , if and only if

$$\limsup_n \sup_{x \in I} |f_n(x) - f(x)| = 0.$$

☐

Exercise 338 Show that a sequence of functions $\{f_n\}$ fails to converge to a function f uniformly on an interval I if and only if there is some positive ε_0 so that a sequence $\{x_k\}$ of points in I and a subsequence $\{f_{n_k}\}$ can be found such that

$$|f_{n_k}(x_k) - f(x_k)| \geq \varepsilon_0.$$

□

Exercise 339 Apply the criterion in the preceding exercise to show that the sequence $f_n(x) = x^n$ does not converge uniformly to zero on $(0, 1)$.

□

Exercise 340 Prove Theorem 3.31.

Answer □

Exercise 341 Verify that the geometric series $\sum_{k=0}^{\infty} x^k$, which converges pointwise on $(-1, 1)$, does not converge uniformly there.

□

Exercise 342 Do the same for the series obtained by differentiating the series in Exercise 341; that is, show that $\sum_{k=1}^{\infty} kx^{k-1}$ converges pointwise but not uniformly on $(-1, 1)$. Show that this series does converge uniformly on every closed interval $[a, b]$ contained in $(-1, 1)$.

□

Exercise 343 Verify that the series

$$\sum_{k=1}^{\infty} \frac{\cos kx}{k^2}$$

converges uniformly on $(-\infty, \infty)$.

□

Exercise 344 If $\{f_n\}$ is a sequence of functions converging uniformly on an interval I to a function f , what conditions on the function g would allow you to conclude that $g \circ f_n$ converges uniformly on I to $g \circ f$?

□

Exercise 345 Prove that the series $\sum_{k=0}^{\infty} \frac{x^k}{k}$ converges uniformly on $[0, b]$ for every $b \in [0, 1)$ but does not converge uniformly on $[0, 1)$.

□

Exercise 346 Prove that if $\sum_{k=1}^{\infty} f_k$ converges uniformly on an interval I , then the sequence of terms $\{f_k\}$ converges uniformly on I .

□

3.6. SEQUENCES AND SERIES OF INTEGRALS

107

Exercise 347 A sequence of functions $\{f_n\}$ is said to be uniformly bounded on an interval $[a, b]$ if there is a number M so that

$$|f_n(x)| \leq M$$

for every n and also for every $x \in [a, b]$. Show that a uniformly convergent sequence $\{f_n\}$ of continuous functions on $[a, b]$ must be uniformly bounded. Show that the same statement would not be true for pointwise convergence. $\square$

Exercise 348 Suppose that $f_n \rightarrow f$ on $(-\infty, +\infty)$. What conditions would allow you to compute that

$$\lim_{n \rightarrow \infty} f_n(x + 1/n) = f(x)?$$

$\square$

Exercise 349 Suppose that $\{f_n\}$ is a sequence of continuous functions on the interval $[0, 1]$ and that you know that $\{f_n\}$ converges uniformly on the set of rational numbers inside $[0, 1]$. Can you conclude that $\{f_n\}$ uniformly on $[0, 1]$? (Would this be true without the continuity assertion?) $\square$

Exercise 350 Prove the following variant of the Weierstrass M-test: Let $\{f_k\}$ and $\{g_k\}$ be sequences of functions on an interval I . Suppose that $|f_k(x)| \leq g_k(x)$ for all k and $x \in I$ and that $\sum_{k=1}^{\infty} g_k$ converges uniformly on I . Then the series $\sum_{k=1}^{\infty} f_k$ converges uniformly on I . $\square$

Exercise 351 Prove the following variant on Theorem 3.34: Let $\{a_k\}$ and $\{b_k\}$ be sequences of functions on an interval I . Suppose that $\sum_{k=1}^{\infty} a_k(x)$ converges uniformly on I . Suppose that $\{b_k\}$ is monotone for each $x \in I$ and uniformly bounded on E . Then the series $\sum_{k=1}^{\infty} a_k b_k$ converges uniformly on I . $\square$

Exercise 352 Prove the following variant on Theorem 3.34: Let $\{a_k\}$ and $\{b_k\}$ be sequences of functions on an interval I . Suppose that there is a number M so that

$$\left| \sum_{k=1}^N a_k(x) \right| \leq M$$

for all $x \in I$ and every integer N . Suppose that

$$\sum_{k=1}^{\infty} |b_k - b_{k+1}|$$

converges uniformly on I and that $b_k \rightarrow 0$ uniformly on I . Then the series $\sum_{k=1}^{\infty} a_k b_k$ converges uniformly on I . $\square$

Exercise 353 Prove the following variant on Abel's test (Theorem 3.34): Let $\{a_k(x)\}$ and $\{b_k(x)\}$ be sequences of functions on an interval I . Suppose that $\sum_{k=1}^{\infty} a_k(x)$ converges uniformly on I . Suppose that the series

$$\sum_{k=1}^{\infty} |b_k(x) - b_{k+1}(x)|$$

has uniformly bounded partial sums on I . Suppose that the sequence of functions $\{b_k(x)\}$ is uniformly bounded on I . Then the series $\sum_{k=1}^{\infty} a_k(x)b_k(x)$ converges uniformly on I . $\square$

Exercise 354 Suppose that $\{f_n(x)\}$ is a sequence of continuous functions on an interval $[a, b]$ converging uniformly to a function f on the open interval (a, b) . If f is also continuous on $[a, b]$, show that the convergence is uniform on $[a, b]$. $\square$

Exercise 355 Suppose that $\{f_n\}$ is a sequence of functions converging uniformly to zero on an interval $[a, b]$. Show that $\lim_{n \rightarrow \infty} f_n(x_n) = 0$ for every convergent sequence $\{x_n\}$ of points in $[a, b]$. Give an example to show that this statement may be false if $f_n \rightarrow 0$ merely pointwise. $\square$

Exercise 356 Suppose that $\{f_n\}$ is a sequence of functions on an interval $[a, b]$ with the property that $\lim_{n \rightarrow \infty} f_n(x_n) = 0$ for every convergent sequence $\{x_n\}$ of points in $[a, b]$. Show that $\{f_n\}$ converges uniformly to zero on $[a, b]$. $\square$

3.6.3 Uniform convergence and integrals

We state our main theorem for continuous functions. We know that bounded, continuous functions are integrable and we have several tools that handle unbounded continuous functions.

Theorem 3.35 (uniform convergence of sequences of continuous functions)

Let $f_1, f_2, f_3, \dots$ be a sequence of functions defined and continuous on an open interval (a, b) . Suppose that $\{f_n\}$ converges uniformly on (a, b) to a function f . Then

1. f is continuous on (a, b) .
2. If each f_n is bounded on the interval (a, b) then so too is f .
3. For each closed, bounded interval $[c, d] \subset (a, b)$

$$\lim_{n \rightarrow \infty} \int_c^d f_n(x) dx = \int_c^d \left\{ \lim_{n \rightarrow \infty} f_n(x) \right\} dx = \int_c^d f(x) dx.$$

4. If each f_n is integrable on the interval $[a, b]$ then so too is f and

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \left\{ \lim_{n \rightarrow \infty} f_n(x) \right\} dx = \int_a^b f(x) dx.$$

We have defined uniform convergence of series in a simple way, merely by requiring that the sequence of partial sums converges uniformly. Thus the Corollary follows immediately from the theorem applied to these partial sums.

Corollary 3.36 (uniform convergence of series of continuous functions) *Let $g_1, g_2, g_3, \dots$ be a sequence of functions defined and continuous on an open interval (a, b) . Suppose that the series $\sum_{k=1}^{\infty} g_k$ converges uniformly on (a, b) to a function f . Then*

1. f is continuous on (a, b) .
2. For each closed, bounded interval $[c, d] \subset (a, b)$

$$\sum_{k=1}^{\infty} \int_c^d g_k(x) dx = \int_c^d \left\{ \sum_{k=1}^{\infty} g_k(x) \right\} dx = \int_c^d f(x) dx.$$

3. If each g_k is integrable on the interval $[a, b]$ then so too is f and

$$\sum_{k=1}^{\infty} \int_a^b g_k(x) dx = \int_a^b \left\{ \sum_{k=1}^{\infty} g_k(x) \right\} dx = \int_a^b f(x) dx.$$

Exercise 357 *To prove Theorem 3.35 and its corollary is just a matter of putting together facts that we already know. Do this.* □

3.6.4 A defect of the calculus integral

In the preceding section we have seen that uniform convergence of continuous functions allows for us to interchange the order of integration and limit to obtain the important formula

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \left\{ \lim_{n \rightarrow \infty} f_n(x) \right\} dx.$$

Is this still true if we drop the assumption that the functions f_n are continuous?

We will prove one very weak theorem and give one counterexample to show that the class of integrable functions in the calculus sense is not closed under uniform limits⁶. We will work on this problem again in Section 3.6.6 but we cannot completely handle the defect. We will remedy this defect of the calculus integral in Chapter 4.

⁶Had we chosen back in Section 2.1.1 to accept sequences of exceptional points rather than finite exceptional sets we would not have had this problem here.

Theorem 3.37 Let $f_1, f_2, f_3, \dots$ be a sequence of functions defined and integrable on a closed, bounded interval $[a, b]$. Suppose that $\{f_n\}$ converges uniformly on $[a, b]$ to a function f . Then, provided we assume that f is integrable on $[a, b]$,

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

Exercise 358 Let

$$g_k(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq 1 - \frac{1}{k} \\ 2^{-k} & \text{if } 1 - \frac{1}{k} < x \leq 1. \end{cases}$$

Show that the series $\sum_{k=2}^{\infty} g_k(x)$ of integrable functions converges uniformly on $[0, 1]$ to a function f that is not integrable in the calculus sense. Answer ☐

Exercise 359 Prove Theorem 3.37. Answer ☐

3.6.5 Uniform limits of continuous derivatives

We saw in Section 3.6.3 that a uniformly convergent sequence (or series) of continuous functions can be integrated term-by-term. As an application of our integration theorem we obtain a theorem on term-by-term differentiation. We write this in a form suggesting that the order of differentiation and limit is being reversed.

Theorem 3.38 Let $\{F_n\}$ be a sequence of uniformly continuous functions on an interval $[a, b]$, suppose that each function has a continuous derivative F'_n on (a, b) , and suppose that

1. The sequence $\{F'_n\}$ of derivatives converges uniformly to a function on (a, b) .
2. The sequence $\{F_n\}$ converges pointwise to a function F .

Then F is differentiable on (a, b) and, for all $a < x < b$,

$$F'(x) = \frac{d}{dx} F(x) = \frac{d}{dx} \lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} \frac{d}{dx} F_n(x) = \lim_{n \rightarrow \infty} F'_n(x).$$

For series, the theorem takes the following form:

Corollary 3.39 *Let $\{G_k\}$ be a sequence of uniformly continuous functions on an interval $[a, b]$, suppose that each function has a continuous derivative G'_k on (a, b) , and suppose that*

1. $F(x) = \sum_{k=1}^{\infty} G_k(x)$ pointwise on $[a, b]$.
2. $\sum_{k=1}^{\infty} G'_k(x)$ converges uniformly on (a, b) .

Then, for all $a < x < b$,

$$F'(x) = \frac{d}{dx} F(x) = \frac{d}{dx} \sum_{k=1}^{\infty} G_k(x) = \sum_{k=1}^{\infty} \frac{d}{dx} G_k(x) = \sum_{k=1}^{\infty} G'_k(x).$$

Exercise 360 Using Theorem 3.35, prove Theorem 3.38.

Answer ☐

Exercise 361 Starting with the geometric series

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \quad \text{on } (-1, 1), \quad (3.5)$$

show how to obtain

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1} \quad \text{on } (-1, 1). \quad (3.6)$$

[Note that the series $\sum_{k=1}^{\infty} kx^{k-1}$ does not converge uniformly on $(-1, 1)$. Is this troublesome?]

Answer ☐

Exercise 362 Starting with the definition

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad \text{on } (-\infty, \infty), \quad (3.7)$$

show how to obtain

$$\frac{d}{dx} e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x \quad \text{on } (-\infty, \infty). \quad (3.8)$$

[Note that the series $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ does not converge uniformly on $(-\infty, \infty)$. Is this troublesome?]

Answer ☐

3.6. SEQUENCES AND SERIES OF INTEGRALS

113

Exercise 363 Can the sequence of functions $f_n(x) = \frac{\sin nx}{n^3}$ be differentiated term-by-term? □

Exercise 364 Can the series of functions $\sum_{k=1}^{\infty} \frac{\sin kx}{k^3}$ be differentiated term-by-term? □

Exercise 365 Verify that the function

$$y(x) = 1 + \frac{x^2}{1!} + \frac{x^4}{2!} + \frac{x^6}{3!} + \frac{x^8}{4!} + \dots$$

is a solution of the differential equation $y' = 2xy$ on $(-\infty, \infty)$ without first finding an explicit formula for $y(x)$. □

3.6.6 Uniform limits of discontinuous derivatives

The following theorem reduces the hypotheses of Theorem 3.38 and, accordingly is much more difficult to prove. Here we have dropped the continuity of the derivatives as an assumption.

Theorem 3.40 Let $\{f_n\}$ be a sequence of uniformly continuous functions defined on an interval $[a, b]$. Suppose that $f'_n(x)$ exists for each n and each $x \in (a, b)$ except possibly for x in some finite set C . Suppose that the sequence $\{f'_n\}$ of derivatives converges uniformly on $(a, b) \setminus C$ and that there exists at least one point $x_0 \in [a, b]$ such that the sequence of numbers $\{f_n(x_0)\}$ converges. Then the sequence $\{f_n\}$ converges uniformly to a function f on the interval $[a, b]$, f is differentiable with, at each point $x \in (a, b) \setminus C$,

$$f'(x) = \lim_{n \rightarrow \infty} f'_n(x) \quad \text{and} \quad \lim_{n \rightarrow \infty} \int_a^b f'_n(x) dx = \int_a^b f'(x) dx.$$

Exercise 366 Prove Theorem 3.40.

Answer □

Exercise 367 For infinite series, how can Theorem 3.40 be rewritten?

Answer □

Exercise 368 (uniform limits of integrable functions) At first sight Theorem 3.40 seems to supply the following observation: If $\{g_n\}$ is a sequence of functions integrable in the calculus sense on an interval $[a, b]$ and g_n converges uniformly to a function g on $[a, b]$ then g must also be integrable. Is this true? Answer $\square$

Exercise 369 In the statement of Theorem 3.40 we hypothesized the existence of a single point x_0 at which the sequence $\{f_n(x_0)\}$ converges. It then followed that the sequence $\{f_n\}$ converges on all of the interval I . If we drop that requirement but retain the requirement that the sequence $\{f'_n\}$ converges uniformly to a function g on I , show that we cannot conclude that $\{f_n\}$ converges on I , but we can still conclude that there exists f such that $f' = g = \lim_{n \rightarrow \infty} f'_n$ on I . Answer $\square$

3.7 The monotone convergence theorem

Two of the most important computations with integrals are taking a limit inside an integral,

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \left(\lim_{n \rightarrow \infty} f_n(x) \right) dx$$

and summing a series inside an integral,

$$\sum_{k=1}^{\infty} \int_a^b g_k(x) dx = \int_a^b \left(\sum_{k=1}^{\infty} g_k(x) \right) dx.$$

The counterexamples in Section 3.6.1, however, have made us very wary of doing this. The uniform convergence results of Section 3.6.5, on the other hand, have encouraged us to check for uniform convergence as a guarantee that these operations will be successful.

But uniform convergence is not a *necessary* requirement. There are important weaker assumptions that will allow us to use sequence and series techniques on integrals. For sequences an assumption that the sequence is monotone will work. For series an assumption that the terms are nonnegative will work.

3.7.1 Summing inside the integral

We establish that the summation formula

$$\int_a^b \left(\sum_{n=1}^{\infty} g_k(x) \right) dx = \sum_{n=1}^{\infty} \left(\int_a^b g_k(x) dx \right)$$

is possible for *nonnegative* functions. We need also to assume that the sum function $f(x) = \sum_{n=1}^{\infty} g_k(x)$ is itself integrable since that cannot be deduced otherwise.

This is just a defect in the calculus integral; in a more general theory of integration we would be able to conclude both that the sum is indeed integrable and also that the sum formula is correct. (See Part Two of this text.) This defect is more serious than it might appear. In most applications the only thing we might know about the function $f(x) = \sum_{n=1}^{\infty} g_k(x)$ is that it is the sum of this series. We may not be able to check continuity and we certainly are unlikely to be able to find an indefinite integral.

We split the statement into two lemmas for ease of proof. Together they supply the integration formula for the sum of nonnegative integrable functions.

Lemma 3.41 Suppose that $f, g_1, g_2, g_3, \dots$ is a sequence of nonnegative functions, each one integrable on a closed bounded interval $[a, b]$. If, for all but finitely many x in (a, b)

$$f(x) \geq \sum_{k=1}^{\infty} g_k(x),$$

then

$$\int_a^b f(x) dx \geq \sum_{k=1}^{\infty} \left(\int_a^b g_k(x) dx \right). \quad (3.9)$$

Lemma 3.42 Suppose that $f, g_1, g_2, g_3, \dots$ is a sequence of nonnegative functions, each one integrable on a closed bounded interval $[a, b]$. If, for all but finitely many x in (a, b) ,

$$f(x) \leq \sum_{k=1}^{\infty} g_k(x),$$

then

$$\int_a^b f(x) dx \leq \sum_{k=1}^{\infty} \left(\int_a^b g_k(x) dx \right). \quad (3.10)$$

Exercise 370 In each of the lemmas show that we may assume, without loss of generality, that the inequalities

$$f(x) \leq \sum_{k=1}^{\infty} g_k(x), \quad \text{or} \quad f(x) \geq \sum_{k=1}^{\infty} g_k(x),$$

hold for all values of x in the entire interval $[a, b]$.

Answer ☐

Exercise 371 Prove the easier of the two lemmas.

Answer ☐

Exercise 372 Prove Lemma 3.42, or rather give it a try and then consult the write up in the answer section. This is just an argument manipulating Riemann sums so it is not particularly deep; even so it requires some care.

Answer ☐

Exercise 373 Construct an example of a convergent series of continuous functions that converges pointwise to a function that is not integrable in the calculus sense.

☐

3.7.2 Monotone convergence theorem

The series formula immediately supplies the monotone convergence theorem.

Theorem 3.43 (Monotone convergence theorem) Let $f_n : [a, b] \rightarrow \mathbb{R}$ ($n = 1, 2, 3, \dots$) be a nondecreasing sequence of functions, each integrable on the interval $[a, b]$ and suppose that

$$f(x) = \lim_{n \rightarrow \infty} f_n(x)$$

for every x in $[a, b]$ with possibly finitely many exceptions. Then, provided f is also integrable on $[a, b]$,

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

Exercise 374 Deduce Theorem 3.43 from Lemmas 3.41 and 3.42.

Answer ☐

Exercise 375 Prove Theorem 3.43 directly by a suitable Riemann sums argument.

Answer ☐

Exercise 376 Construct an example of a convergent, monotonic sequence of continuous functions that converges pointwise to a function that is not integrable in the calculus sense.

☐

3.8 Integration of power series

A *power series* is an infinite series of the form

$$f(x) = \sum_{n=0}^{\infty} a_n (x-c)^n = a_0 + a_1(x-c)^1 + a_2(x-c)^2 + a_3(x-c)^3 + \dots$$

where a_n is called the *coefficient* of the n th term and c is a constant. One usually says that the series is *centered* at c . By a simple change of variables any power series can be centered at zero and so all of the theory is usually stated for such a power series

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots$$

The set of points x where the series converges is called the *interval of convergence*. (We could call it a *set* of convergence, but we are anticipating that it will turn out to be an interval.)

The main concern we shall have in this chapter is the integration of such series. The topic of power series in general is huge and central to much of mathematics. We can present a fairly narrow picture but one that is complete only insofar as applications of integration theory are concerned.

Theorem 3.44 (convergence of power series) *Let*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots$$

be a power series. Then there is a number R , $0 \leq R \leq \infty$, called the radius of convergence of the series, so that

1. *If $R = 0$ then the series converges only for $x = 0$.*
2. *If $R > 0$ the series converges absolutely for all x in the interval $(-R, R)$.*
3. *If $0 < R < \infty$ the interval of convergence for the series is one of the intervals*

$$(-R, R), (-R, R], [-R, R) \text{ or } [-R, R]$$

and at the endpoints the series may converge absolutely or nonabsolutely.

The next theorem establishes the continuity of a power series within its interval of convergence.

Theorem 3.45 (continuity of power series) *Let*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots$$

be a power series with a radius of convergence R , $0 < R \leq \infty$. Then

1. *f is a continuous function on its interval of convergence [i.e., continuous at all interior points and continuous on the right or left at an endpoint if that endpoint is included].*
2. *If $0 < R < \infty$ and the interval of convergence for the series is $[-R, R]$ then f is uniformly continuous on $[-R, R]$.*

Finally we are in position to show that term-by-term integration of power series is possible in nearly all situations.

3.8. INTEGRATION OF POWER SERIES

Theorem 3.46 (integration of power series) *Let*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots$$

be a power series and let

$$F(x) = \sum_{n=0}^{\infty} \frac{a_n x^{n+1}}{n+1} = a_0 x + a_1 x^2/2 + a_2 x^3/3 + a_3 x^4/4 + \dots$$

be its formally integrated series. Then

1. *Both series have the same radius of convergence R , but not necessarily the same interval of convergence.*
2. *If $R > 0$ then $F'(x) = f(x)$ for every x in $(-R, R)$ and so F is an indefinite integral for f on the interval $(-R, R)$.*
3. *f is integrable on any closed, bounded interval $[a, b] \subset (-R, R)$ and*

$$\int_a^b f(x) dx = F(b) - F(a).$$

4. *If the interval of convergence of the integrated series for F is $[-R, R]$ then f is integrable on $[-R, R]$ and*

$$\int_{-R}^R f(x) dx = F(R) - F(-R).$$

5. *If the interval of convergence of the integrated series for F is $(-R, R]$ then f is integrable on $[0, R]$ and*

$$\int_0^R f(x) dx = F(R) - F(0).$$

6. *If the interval of convergence of the integrated series for F is $[-R, R)$ then f is integrable on $[-R, 0]$ and*

$$\int_{-R}^0 f(x) dx = F(0) - F(-R).$$

Note that the integration theorem uses the interval of convergence of the integrated series. It is not a concern whether the original series for f converges at the endpoints of the interval of convergence, but it is essential to look at these endpoints for the integrated series. The proofs of the separate statements in the two theorems appear in various of the exercises. Note that, while we are interested in integration problems here the proofs are all about derivatives; this is not surprising since the calculus integral itself is simply about derivatives.

Exercise 377 Compute, if possible, the integrals

$$\int_0^1 \left(\sum_{n=0}^{\infty} x^n \right) dx \quad \text{and} \quad \int_{-1}^0 \left(\sum_{n=0}^{\infty} x^n \right) dx.$$

Answer ☐

Exercise 378 Repeat the previous exercise but use only the fact that

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + x^4 + \cdots = \frac{1}{1-x}.$$

Is the answer the same?

Answer ☐

Exercise 379 (careless student) “But,” says the careless student, “both of Exercises 377 and 378 are wrong surely. After all, the series

$$f(x) = 1 + x + x^2 + x^3 + x^4 + \cdots +$$

converges only on the interval $(-1, 1)$ and diverges at the endpoints $x = 1$ and $x = -1$ since

$$1 - 1 + 1 - 1 + 1 - 1 + 1 - 1 = ?$$

and

$$1 + 1 + 1 + 1 + \cdots = \infty.$$

You cannot expect to integrate on either of the intervals $[-1, 0]$ or $[0, 1]$.” What is your response?

Answer ☐

Exercise 380 (calculus student notation) For most calculus students it is tempting to write

$$\int (a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots) dx = \int a_0 dx + \int a_1x dx + \int a_2x^2 dx + \int a_3x^3 dx + \cdots$$

3.8. INTEGRATION OF POWER SERIES

121

Is this a legitimate interpretation of this indefinite integral?

Answer ☐

Exercise 381 (calculus student notation) For most calculus students it is tempting to write

$$\int_a^b (a_0 + a_1x + a_2x^2 + a_3x^3 + \dots) dx = \int_a^b a_0 dx + \int_a^b a_1x dx + \int_a^b a_2x^2 dx + \int_a^b a_3x^3 dx + \dots$$

Is this a legitimate interpretation of this definite integral?

Answer ☐

Exercise 382 Show that the series

$$f(x) = 1 + 2x + 3x^2 + 4x^3 + \dots$$

has a radius of convergence 1 and an interval of convergence exactly equal to $(-1, 1)$. Show that f is not integrable on $[0, 1]$, but that it is integrable $[-1, 0]$ and yet the computation

$$\begin{aligned} \int_{-1}^0 (1 + 2x + 3x^2 + 4x^3 + \dots) dx &= \int_{-1}^0 dx + \int_{-1}^0 2x dx + \int_{-1}^0 3x^2 dx + \int_{-1}^0 4x^3 dx + \dots \\ &= -1 + 1 - 1 + 1 - 1 + 1 - \dots \end{aligned}$$

cannot be used to evaluate the integral.

Note: Since the interval of convergence of the integrated series is also $(-1, 1)$, Theorem 3.46 has nothing to say about whether f is integrable on $[0, 1]$ or $[-1, 0]$.

Answer ☐

Exercise 383 Determine the radius of convergence of the series

$$\sum_{k=1}^{\infty} k^k x^k = x + 4x^2 + 27x^3 + \dots$$

Answer ☐

Exercise 384 Show that, for every $0 \leq s \leq \infty$, there is a power series whose radius of convergence R is exactly s .

Answer ☐

Exercise 385 Show that the radius of convergence of a series

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

can be described as

$$R = \sup\{r : 0 < r \text{ and } \sum_{k=0}^{\infty} a_k r^k \text{ converges}\}.$$

□

Exercise 386 (root test for power series) Show that the radius of convergence of a series

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

is given by the formula

$$R = \frac{1}{\limsup_{k \rightarrow \infty} \sqrt[k]{|a_k|}}.$$

□

Exercise 387 Show that the radius of convergence of the series

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

is the same as the radius of convergence of the formally differentiated series

$$a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + \dots$$

□

Exercise 388 Show that the radius of convergence of the series

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

is the same as the radius of convergence of the formally integrated series

$$a_0x + a_1x^2/2 + a_2x^3/3 + a_3x^4/4 + \dots$$

Answer □

Exercise 389 (ratio test for power series) Show that the radius of convergence of the series

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

3.8. INTEGRATION OF POWER SERIES

123

is given by the formula

$$R = \lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right|,$$

assuming that this limit exists or equals ∞ .

Answer ☐

Exercise 390 (ratio/root test for power series) Give an example of a power series for which the radius of convergence R satisfies

$$R = \frac{1}{\lim_{k \rightarrow \infty} \sqrt[k]{|a_k|}}$$

but

$$\lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right|$$

does not exist.

Answer ☐

Exercise 391 (ratio test for power series) Give an example of a power series for which the radius of convergence R satisfies

$$\liminf_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| < R < \limsup_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right|.$$

Note: for such a series the ratio test cannot give a satisfactory estimate of the radius of convergence.

Answer ☐

Exercise 392 If the coefficients $\{a_k\}$ of a power series

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

form a bounded sequence show that the radius of convergence is at least 1.

Answer ☐

Exercise 393 If the coefficients $\{a_k\}$ of a power series

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

form an unbounded sequence show that the radius of convergence is no more than 1.

Answer ☐

Exercise 394 If the power series

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

has a radius of convergence R_a and the power series

$$b_0 + b_1x + b_2x^2 + b_3x^3 + \dots$$

has a radius of convergence R_b and $|a_k| \leq |b_k|$ for all k sufficiently large, what relation must hold between R_a and R_b ?

Answer $\square$

Exercise 395 If the power series

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

has a radius of convergence R , what must be the radius of convergence of the series

$$a_0 + a_1x^2 + a_2x^4 + a_3x^6 + \dots$$

Answer $\square$

Exercise 396 Suppose that the series

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

has a finite radius of convergence R and suppose that $|x_0| > R$. Show that, not only does

$$a_0 + a_1x_0 + a_2x_0^2 + a_3x_0^3 + \dots$$

diverge but that $\lim_{n \rightarrow \infty} |a_n x_0^n| = \infty$.

$\square$

Exercise 397 Suppose that the series

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

has a positive radius of convergence R . Use the Weierstrass M-test to show that the series converges uniformly on any closed, bounded subinterval $[a, b] \subset (-R, R)$.

$\square$

Exercise 398 Suppose that the series

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

has a positive radius of convergence R . Use Exercise 397 to show that f is differentiable on $(-R, R)$ and that, for all x in that interval,

$$f'(x) = a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + \dots$$

$\square$

Exercise 399 Suppose that the series

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

has a positive radius of convergence R . Use Exercise 398 to show that f has an indefinite integral on $(-R, R)$ given by the function

$$F(x) = a_0x + a_1x^2/2 + a_2x^3/3 + a_3x^4/4 + \dots$$

□

Exercise 400 Suppose that the series

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

has a positive, finite radius of convergence R and that the series converges absolutely at one of the two endpoints R or $-R$ of the interval of convergence. Use the Weierstrass M-test to show that the series converges uniformly on $[-R, R]$. Deduce from this that f' is integrable on $[-R, R]$.

Note: this is the best that the Weierstrass M-test can do applied to power series. If the series converges nonabsolutely at one of the two endpoints R or $-R$ of the interval then the test does not help. Answer □

Exercise 401 Suppose that the series

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

has a positive, finite radius of convergence R and that the series converges nonabsolutely at one of the two endpoints R or $-R$ of the interval of convergence. Use a variant of the Abel test for uniform convergence to show that the series converges uniformly on any closed subinterval $[a, b]$ of the interval of convergence. Deduce from this that f' is integrable on any such interval $[a, b]$.

Note: this completes the picture for the integrability problem of this section. Answer □

Exercise 402 What power series will converge uniformly on $(-\infty, \infty)$? □

Exercise 403 Show that if $\sum_{k=0}^{\infty} a_k x^k$ converges uniformly on an interval $(-r, r)$, then it must in fact converge uniformly on $[-r, r]$. Deduce that if the interval of convergence is exactly of the form $(-R, R)$, or $[-R, R)$ or $(-R, R]$, then the series cannot converge uniformly on the entire interval of convergence. Answer □

Exercise 404 Suppose that a function $f(x)$ has two power series representations

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

and

$$f(x) = b_0 + b_1x + b_2x^2 + b_3x^3 + \dots$$

both valid at least in some interval $(-r, r)$ for $r > 0$. What can you conclude? □

Exercise 405 Suppose that a function $f(x)$ has a power series representations

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

valid at least in some interval $(-r, r)$ for $r > 0$. Show that, for each $k = 0, 1, 2, 3, \dots$,

$$a_k = \frac{f^{(k)}(0)}{k!}.$$

□

Exercise 406 In view of Exercise 405 it would seem that we must have the formula

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$$

provided only that the function f is infinitely often differentiable at $x = 0$. Is this a correct observation? Answer □

3.9 Applications of the integral

It would be presumptuous to try to teach here applications of the integral, since those applications are nearly unlimited. But here are a few that follow a simple theme and are traditionally taught in all calculus courses.

The theme takes advantage of the fact that an integral can (under certain hypotheses) be approximated by a Riemann sum

$$\int_a^b f(x) dx \approx \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}).$$

If there is an application where some concept can be expressed as a limiting version of sums of this type, then that concept can be captured by an integral. Whatever the concept is, it must be necessarily “additive” and expressible as

sums of products that can be interpreted as

$$f(\xi_i) \times (x_i - x_{i-1}).$$

The simplest illustration is area. We normally think of area as additive. We can interpret the product

$$f(\xi_i) \times (x_i - x_{i-1}).$$

as the area of a rectangle with length $(x_i - x_{i-1})$ and height $f(\xi_i)$. The Riemann sum itself then is a sum of areas of rectangles. If we can determine that the area of some figure is approximated by such a sum, then the area can be described completely by an integral.

For applications in physics one might use t as a time variable and then interpret

$$\int_a^b f(t) dt \approx \sum_{i=1}^n f(\tau_i)(t_i - t_{i-1})$$

thinking of $f(\tau_i)$ as some measurement (e.g., velocity, acceleration, force) that is occurring throughout the time interval $[t_{i-1}, t_i]$.

An accumulation point of view For many applications of the calculus the Riemann sum approach is an attractive way of expressing the concepts that arise as a definite integral. There is another way which bypasses Riemann sums and goes directly back to the definition of the integral as an antiderivative.

We can write this method using the slogan

$$\int_a^{x+h} f(t) dt - \int_a^x f(t) dt \approx f(\xi) \times h. \quad (3.11)$$

Suppose that a concept we are trying to measure can be captured by a function $A(x)$ on some interval $[a, b]$. We suppose that we have already measured $A(x)$ and now wish to add on a bit more to get to $A(x+h)$ where h is small. We imagine the new amount that we must add on can be expressed as

$$f(\xi) \times h$$

thinking of $f(\xi)$ as some measurement that is occurring throughout the interval $[x, x+h]$. In that case our model for the concept is the integral $\int_a^b f(t) dt$. This is because (3.11) suggests that $A'(x) = f(x)$.

3.9.1 Area and the method of exhaustion

There is a long historical and cultural connection between the theory of integration and the geometrical theory of area. Usually one takes the following as the primary definition of area.

Definition 3.47 Let $f : [a, b] \rightarrow \mathbb{R}$ be an integrable, nonnegative function and suppose that $R(f, a, b)$ denotes the region in the plane bounded on the left by the line $x = a$, on the right by the line $x = b$, on the bottom by the line $y = 0$ and on the top by the graph of the function f (i.e., by $y = f(x)$). Then this region is said to have an area and value of that area is assigned to be

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

The region can also be described by writing it as a set of points:

$$R(f, a, b) = \{(x, y) : a \leq x \leq b, 0 \leq y \leq f(x)\}.$$

We can justify this definition by the method of Riemann sums combined with a method of the ancient Greeks known as the method of exhaustion of areas.

Let us suppose that $f : [a, b] \rightarrow \mathbb{R}$ is a uniformly continuous, nonnegative function and suppose that $R(f, a, b)$ is the region as described above. Take any subdivision

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

Then there must exist points $\xi_i, \eta_i \in [x_{i-1}, x_i]$ for $i = 1, 2, \dots, n$ so that $f(\xi_i)$ is the maximum value of f in the interval $[x_{i-1}, x_i]$ and $f(\eta_i)$ is the minimum value of f in that interval. We consider the two partitions

$$\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\} \quad \text{and} \quad \{([x_i, x_{i-1}], \eta_i) : i = 1, 2, \dots, n\}$$

and the two corresponding Riemann sums

$$\sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \quad \text{and} \quad \sum_{i=1}^n f(\eta_i)(x_i - x_{i-1}).$$

The larger sum is greater than the integral $\int_a^b f(x) dx$ and the smaller sum is lesser than that number. This is because there is a choice of points ξ_i^* that is exactly equal to the integral,

$$\int_a^b f(x) dx = \sum_{i=1}^n f(\xi_i^*)(x_i - x_{i-1})$$

and here we have $f(\eta_i) \leq f(\xi_i^*) \leq f(\xi_i)$. (See Section 3.3.1.)

But if the region were to have an “area” we would expect that area is also between these two sums. That is because the larger sum represents the area of a collection of n rectangles that include our region and the smaller sum represents the area of a collection of n rectangles that are included inside our region. If we consider all possible subdivisions then the same situation holds: the area of the region (if it has one) must lie between the upper sums and the lower sums. But according to Theorem 3.17 the only number with this property is the integral $\int_a^b f(x) dx$ itself.

Certainly then, for continuous functions anyway, this definition of the area of such a region would be compatible with any other theory of area.

Exercise 407 (an accumulation argument) Here is another way to argue that integration theory and area theory must be closely related. Imagine that area has some (at the moment) vague meaning to you. Let $f : [a, b] \rightarrow \mathbb{R}$ be a uniformly continuous, nonnegative function. For any $a \leq s < t \leq b$ let $A(f, s, t)$ denote the area of the region in the plane bounded on the left by the line $x = s$, on the right by the line $x = t$, on the bottom by the line $y = 0$ and on the top by the curve $y = f(x)$. Argue for each of the following statements:

1. $A(f, a, s) + A(f, s, t) = A(f, a, t)$.
2. If $m \leq f(x) \leq M$ for all $s \leq x \leq t$ then $m(t - s) \leq A(f, s, t) \leq M(t - s)$.
3. At any point $a < x < b$,

$$\frac{d}{dx} A(f, a, x) = f(x).$$

4. At any point $a < x < b$,

$$A(f, a, x) = \int_a^x f(t) dt.$$

Answer $\square$

Exercise 408 Show that the area of the triangle

$$\{(x, y) : a \leq x \leq b, 0 \leq y \leq m(x - a)\}.$$

is exactly as you would normally have computed it precalculus. □

Exercise 409 Show that the area of the trapezium

$$\{(x, y) : a \leq x \leq b, 0 \leq y \leq c + m(x - a)\}.$$

is exactly as you would normally have computed it precalculus. □

Exercise 410 Show that the area of the half-circle

$$\{(x, y) : -1 \leq x \leq 1, 0 \leq y \leq \sqrt{1 - x^2}\}.$$

is exactly as you would normally have computed it precalculus. Answer □

Exercise 411 One usually takes this definition for the area between two curves:

Definition 3.48 Let $f, g : [a, b] \rightarrow \mathbb{R}$ be integrable functions and suppose that $f(x) \geq g(x)$ for all $a \leq x \leq b$. Let $R(f, g, a, b)$ denote the region in the plane bounded on the left by the line $x = a$, on the right by the line $x = b$, on the bottom by the curve $y = g(x)$ and on the top by the curve by $y = f(x)$. Then this region is said to have an area and value of that area is assigned to be

$$\int_a^b [f(x) - g(x)] dx.$$

Use this definition to find the area inside the circle $x^2 + y^2 = r^2$. Answer □

Exercise 412 Using Definition 3.48 compute the area between the graphs of the functions $g(x) = 1 + x^2$ and $h(x) = 2x^2$ on $[0, 1]$. Explain why the Riemann sum

$$\sum_{i=1}^n [g(\xi_i) - h(\xi_i)](x_i - x_{i-1})$$

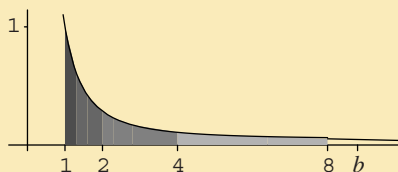


Figure 3.4: Computation of an area by $\int_1^\infty x^{-2} dx$.

and the corresponding integral $\int_0^1 [g(x) - h(x)] dx$ cannot be interpreted using the method of exhaustion to be computing both upper and lower bounds for this area. Discuss. Answer ☐

Exercise 413 In Figure 3.4 we show graphically how to interpret the area that is represented by $\int_1^\infty x^{-2} dx$. Note that

$$\int_1^2 x^{-2} dx = 1/2, \int_2^4 x^{-2} dx = 1/4, \int_4^8 x^{-2} dx = 1/8$$

and so we would expect

$$\int_1^\infty x^{-2} dx = 1/2 + 1/4 + 1/8 + \dots$$

Check that this is true. Answer ☐

3.9.2 Volume

A full treatment of the problem of defining and calculating volumes is outside the scope of a calculus course that focusses only on integrals of this type:

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

But if the problem addresses a very special type of volume, those volumes obtained by rotating a curve about some line, then often the formula

$$\pi \int_a^b [f(x)]^2 dx$$

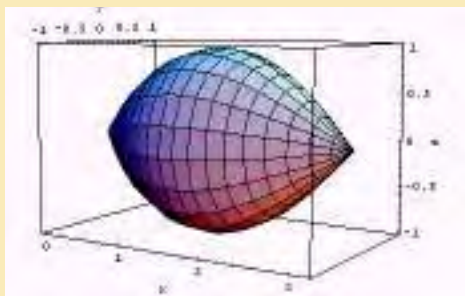


Figure 3.5: $\sin x$ rotated around the x -axis.

can be interpreted as providing the correct volume interpretation and computation.

Once again the justification is the method of exhaustion. We assume that volumes, like areas, are additive. We assume that a correct computation of the volume of cylinder that has radius r and height h is $\pi r^2 h$. In particular the volume of a cylinder that has radius $f(\xi_i)$ and height $(x_i - x_{i-1})$ is

$$\pi [f(\xi_i)]^2 (x_i - x_{i-1}).$$

The total volume for a collection of such cylinders would be (since we assume volume is additive)

$$\pi \sum_{i=1}^n [f(\xi_i)]^2 (x_i - x_{i-1}).$$

We then have a connection with the formula

$$\pi \int_a^b [f(x)]^2 dx.$$

One example with suitable pictures illustrates the method. Take the graph of the function $f(x) = \sin x$ on the interval $[0, \pi]$ and rotate it (into three dimensional space) around the x -axis. Figure 3.5 shows the football (i.e., American football) shaped object.

Subdivide the interval $[0, \pi]$,

$$0 = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = \pi$$

Then there must exist points $\xi_i, \eta_i \in [x_{i-1}, x_i]$ for $i = 1, 2, \dots, n$ so that $\sin(\xi_i)$ is the maximum value of $\sin x$ in the interval

$[x_{i-1}, x_i]$ and $\sin(\eta_i)$ is the minimum value of $\sin x$ in that interval. We consider the two partitions

$$\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\} \quad \text{and} \quad \{([x_i, x_{i-1}], \eta_i) : i = 1, 2, \dots, n\}$$

and the two corresponding Riemann sums

$$\pi \sum_{i=1}^n \sin^2(\xi_i)(x_i - x_{i-1}) \quad \text{and} \quad \pi \sum_{i=1}^n \sin^2(\eta_i)(x_i - x_{i-1}).$$

The “football” is entirely contained inside the cylinders representing the first sum and the cylinders representing the second sum are entirely inside the football.

There is only one value that lies between these sums for all possible choice of partition, namely the number

$$\pi \int_0^\pi [\sin x]^2 dx.$$

We know this because this integral can be uniformly approximated by Riemann sums. The method of exhaustion then claims that the volume of the football must be this number.

In general this argument justifies the following working definition. This is the analogue for volumes of revolution of Definition 3.48 .

Definition 3.49 *Let f and g be continuous, nonnegative functions on an interval $[a, b]$ and suppose that $g(x) \leq f(x)$ for all $a \leq x \leq b$. Then the volume of the solid obtained by rotating the region between the two curves $y = f(x)$ and $y = g(x)$ about the x -axis is given by*

$$\pi \int_a^b ([f(x)]^2 - [g(x)]^2) dx.$$

Exercise 414 (shell method) *There is a similar formula for a volume of revolution when the curve $y = f(x)$ on $[a, b]$ (with $a < 0$) is rotated about the y -axis. One can either readjust by interchanging x and y to get a formula of the form $\pi \int_c^d [g(y)]^2 dy$ or use the so-called shell method that has a formula*

$$2\pi \int_a^b x \times h dx$$

where h is a height measurement in the shell method. Investigate. □

Exercise 415 (surface area) *If a nonnegative function $y = f(x)$ is continuously differentiable throughout the interval*

$[a, b]$, then the formula for the area of the surface generated by revolving the curve about the x -axis is generally claimed to be

$$\pi \int_a^b 2f(x) \sqrt{1 + [f'(x)]^2} dx.$$

Using the same football studied in this section how could you justify this formula. □

3.9.3 Length of a curve

In mathematics a *curve* [sometimes called a *parametric curve*] is a pair of uniformly continuous functions F, G defined on an interval $[a, b]$. The points $(F(t), G(t))$ in the plane are considered to trace out the curve as t moves from the endpoint a to the endpoint b . The curve is thought of as a mapping taking points in the interval $[a, b]$ to corresponding points in the plane. Elementary courses often express the curve this way,

$$x = F(t), y = G(t) \quad a \leq t \leq b,$$

referring to the two equations as *parametric equations* for the curve and to the variable t as a *parameter*.

The set of points

$$\{(x, y) : x = F(t), y = G(t), a \leq t \leq b\}$$

is called the *graph of the curve*. It is *not* the curve itself but, for novices, it may be difficult to make this distinction. The curve is thought to be *oriented* in the sense that as t moves in its positive direction [i.e., from a to b] the curve is traced out in that order. Any point on the curve may be covered many times by the curve itself; the curve can cross itself or be very complicated indeed, even though the graph might be simple.

For example, take any continuous function F on $[0, 1]$ with $F(0) = 0$ and $F(1) = 1$ and $0 \leq F(x) \leq 1$ for $0 \leq x \leq 1$. Then the curve $(F(t), F(t))$ traces out the points on the line connecting $(0, 0)$ to $(1, 1)$. But the points can be traced and retraced many times and the “trip” itself may have infinite length. All this even though the line segment itself is simple and short (it has length $\sqrt{2}$).

The length of a curve is defined by estimating the length of the route taken by the curve by approximating its length by a polygonal path. Subdivide the interval

$$a = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = b$$

and then just compute the length of a trip to visit each of the points $(F(a), G(a)), (F(t_1), G(t_1)), (F(t_2), G(t_2)), \dots, (F(b), G(b))$ in that order. The definition should resemble our definition of a function of bounded variation and, indeed,

the two ideas are very closely related.

Definition 3.50 (rectifiable curve) A curve given by a pair of functions $F, G : [a, b] \rightarrow \mathbb{R}$ is said to be rectifiable if there is a number M so that

$$\sum_{i=1}^n \sqrt{[F(t_i) - F(t_{i-1})]^2 + [G(t_i) - G(t_{i-1})]^2} \leq M$$

for all choices of points

$$a = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = b.$$

The least such number M is called the length of the curve.

Exercise 416 Show that a curve given by a pair of uniformly continuous functions $F, G : [a, b] \rightarrow \mathbb{R}$ is rectifiable if and only if both functions F and G have bounded variation on $[a, b]$. Obtain, moreover, that the length L of the curve must satisfy

$$\max\{V(F, [a, b]), V(G, [a, b])\} \leq L \leq V(F, [a, b]) + V(G, [a, b]).$$

Answer ☐

Exercise 417 Prove the following theorem which supplies the familiar integral formula for the length of a curve.

Theorem 3.51 Suppose that a curve is given by a pair of uniformly continuous functions $F, G : [a, b] \rightarrow \mathbb{R}$ and suppose that both F and G have bounded, continuous derivatives at every point of (a, b) with possibly finitely many exceptions. Then the curve is rectifiable and, moreover, the length L of the curve must satisfy

$$L = \int_a^b \sqrt{[F'(t)]^2 + [G'(t)]^2} dt.$$

Answer ☐

Exercise 418 Take any continuous function F on $[0, 1]$ with $F(0) = 0$ and $F(1) = 1$ and $0 \leq F(x) \leq 1$ for $0 \leq x \leq 1$. Then the curve $(F(t), F(t))$ traces out the points on the line segment connecting $(0, 0)$ to $(1, 1)$. Why does the graph of the curve contain all points on the line segment? Answer $\square$

Exercise 419 Find an example of a continuous function F on $[0, 1]$ with $F(0) = 0$ and $F(1) = 1$ and $0 \leq F(x) \leq 1$ for $0 \leq x \leq 1$ such that the curve $(F(t), F(t))$ has infinite length. Can you find an example where the length is 2? Can you find one where the length is 1? Which choices will have length equal to $\sqrt{2}$ which is, after all, the actual length of the graph of the curve? $\square$

Exercise 420 A curve in three dimensional space is a triple of uniformly continuous functions $(F(t), G(t), H(t))$ defined on an interval $[a, b]$. Generalize to the theory of such curves the notions presented in this section for curves in the plane. $\square$

Exercise 421 The graph of a uniformly continuous function $f : [a, b] \rightarrow \mathbb{R}$ may be considered a curve in this sense using the pair of functions $F(t) = t$, $G(t) = f(t)$ for $a \leq t \leq b$. This curve has for its graph precisely the graph of the function, i.e., the set

$$\{(x, y) : y = f(x) \text{ } a \leq x \leq b\}.$$

Under this interpretation the graph of the function has a length if this curve has a length. Discuss. Answer $\square$

Exercise 422 Find the length of the graph of the function

$$f(x) = \frac{1}{2}(e^x + e^{-x}), \quad 0 \leq x \leq 2.$$

[The answer is $\frac{1}{2}(e^2 - e^{-2})$. This is a typical question in a calculus course, chosen not because the curve is of great interest, but because it is one of the very few examples that can be computed by hand.] Answer $\square$

3.10 Numerical methods

This is a big subject with many ideas and many pitfalls. As a calculus student you are mainly [but check with your instructor] responsible for learning a few standard methods, eg., the trapezoidal rule and Simpson's rule.

In any practical situation where numbers are needed how might we compute

$$\int_a^b f(x) dx?$$

The computation of any integral would seem (judging by the definition) to require first obtaining an indefinite integral F [checking to see it is continuous, of course, and that $F'(x) = f(x)$ at all but finitely many points in (a, b)]. Then the formula

$$\int_a^b f(x) dx = F(b) - F(a)$$

would give the precise value.

But finding an indefinite integral may be impractical. There must be an indefinite integral if the integral exists, but that does not mean that it must be given by an accessible formula or that we would have the skills to find it. The history of our subject is very long so many problems have already been solved but finding antiderivatives is most often not the best method even when it is possible to carry it out.

Finding a close enough value for $\int_a^b f(x) dx$ may be considerably easier and less time consuming than finding an indefinite integral. The former is just a number, the latter is a function, possibly mysterious.

Just use Riemann sums? If we have no knowledge whatever about the function f beyond the fact that it is bounded and continuous mostly everywhere then to estimate $\int_a^b f(x) dx$ we could simply use Riemann sums. Divide the interval $[a, b]$ into pieces of equal length h

$$a < a + h < a + 2h < a + 3h < \dots < a + (n-1)h < b.$$

Here there are $n-1$ pieces of equal length and the last piece, the n th piece, has (perhaps) smaller length

$$b - (a + (n-1)h) \leq h.$$

Then

$$\int_a^b f(x) dx \approx [f(\xi_1) + f(\xi_2) + \dots + f(\xi_{n-1})]h + f(\xi_n)[b - (a + (n-1)h)].$$

We do know that, for small enough h , the approximation is as close as we please to the actual value. And we can estimate the error if we know the oscillation of the function in each of these intervals.

If we were to use this in practise then the computation is simpler if we choose always ξ_i as an endpoint of the

corresponding interval and we choose for h only lengths $(b-a)/n$ so that *all* the pieces have equal length. The methods that follow are better for functions that arise in real applications, but if we want a method that works for all continuous functions, there is no guarantee that any other method would surpass this very naive method.

Trapezoidal rule Here is the (current) Wikipedia statement of the rule:

In mathematics, the trapezoidal rule (also known as the trapezoid rule, or the trapezium rule in British English) is a way to approximately calculate the definite integral

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

The trapezoidal rule works by approximating the region under the graph of the function $f(x)$ by a trapezoid and calculating its area. It follows that

$$\int_a^b f(x) dx \approx (b-a) \frac{f(a) + f(b)}{2}.$$

To calculate this integral more accurately, one first splits the interval of integration $[a,b]$ into n smaller subintervals, and then applies the trapezoidal rule on each of them. One obtains the composite trapezoidal rule:

$$\int_a^b f(x) dx \approx \frac{b-a}{n} \left[\frac{f(a) + f(b)}{2} + \sum_{k=1}^{n-1} f\left(a + k \frac{b-a}{n}\right) \right].$$

This can alternatively be written as:

$$\int_a^b f(x) dx \approx \frac{b-a}{2n} (f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n))$$

where

$$x_k = a + k \frac{b-a}{n}, \text{ for } k = 0, 1, \dots, n$$

The error of the composite trapezoidal rule is the difference between the value of the integral and the numerical result:

$$\text{error} = \int_a^b f(x) dx - \frac{b-a}{n} \left[\frac{f(a) + f(b)}{2} + \sum_{k=1}^{n-1} f\left(a + k \frac{b-a}{n}\right) \right].$$

This error can be written as

$$\text{error} = -\frac{(b-a)^3}{12n^2} f''(\xi),$$

where ξ is some number between a and b .

It follows that if the integrand is concave up (and thus has a positive second derivative), then the error is negative and the trapezoidal rule overestimates the true value. This can also be seen from the geometric picture: the trapezoids include all of the area under the curve and extend over it. Similarly, a concave-down function yields an underestimate because area is unaccounted for under the curve, but none is counted above. If the interval of the integral being approximated includes an inflection point, then the error is harder to identify.

Simpson's rule Simpson's rule is another method for numerical approximation of definite integrals. The approximation on a single interval uses the endpoints and the midpoint. In place of a trapezoidal approximation, an approximation using quadratics produces:

$$\int_a^b f(x) dx \approx \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right].$$

It is named after the English mathematician Thomas Simpson (1710–1761). An extended version of the rule for $f(x)$ tabulated at $2n$ evenly spaced points a distance h apart,

$$a = x_0 < x_1 < \cdots < x_{2n} = b$$

is

$$\int_{x_0}^{x_{2n}} f(x) dx = \frac{h}{3} [f_0 + 4(f_1 + f_3 + \cdots + f_{2n-1}) + 2(f_2 + f_4 + \cdots + f_{2n-2}) + f_{2n}] - R_n,$$

where $f_i = f(x_i)$ and where the remainder term is

$$R_n = \frac{nh^5 f'''(\xi)}{90}$$

for some $\xi \in [x_0, x_{2n}]$.

Exercise 423 Show that the trapezoidal rule can be interpreted as asserting that a reasonable computation of the mean

value of a function on an interval,

$$\frac{1}{b-a} \int_a^b f(x) dx,$$

is simply to average the values of the function at the two endpoints.

Answer ☐

Exercise 424 Establish the identity

$$\int_a^b f(x) dx = \frac{f(a) + f(b)}{2}(b-a) - \frac{1}{2} \int_a^b (x-a)(b-x) f''(x) dx$$

under suitable hypotheses on f .

Answer ☐

Exercise 425 Establish the identity

$$\int_a^b f(x) dx - \frac{f(a) + f(b)}{2}(b-a) = -\frac{(b-a)^3 f''(\xi)}{12}$$

for some point $a < \xi < b$, under suitable hypotheses on f .

Answer ☐

Exercise 426 Establish the inequality

$$\left| \int_a^b f(x) dx - \frac{f(a) + f(b)}{2}(b-a) \right| \leq \frac{(b-a)^2}{8} \int_a^b |f''(x)| dx.$$

under suitable hypotheses on f .

Answer ☐

Exercise 427 Prove the following theorem and use it to provide the estimate for the error given in the text for an application of the trapezoidal rule.

Theorem 3.52 Suppose that f is twice continuously differentiable at all points of the interval $[a, b]$. Let

$$T_n = \frac{b-a}{n} \left[\frac{f(a) + f(b)}{2} + \sum_{k=1}^{n-1} f\left(a + k \frac{b-a}{n}\right) \right]$$

denote the usual trapezoidal sum for f . Then

$$\int_a^b f(x) dx - T_n = - \sum_{k=1}^n \frac{(b-a)^3}{12n^3} f''(\xi_i)$$

for appropriately chosen points ξ_i in each interval

$$[x_{i-1}, x_i] = \left[a + \frac{(i-1)(b-a)}{n}, a + \frac{i(b-a)}{n} \right] \quad (i = 1, 2, 3, \dots, n)$$

Answer ☐

Exercise 428 Prove the following theorem which elaborates on the error in the trapezoidal rule.

Theorem 3.53 Suppose that f is twice continuously differentiable at all points of the interval $[a, b]$. Let

$$T_n = \frac{b-a}{n} \left[\frac{f(a) + f(b)}{2} + \sum_{k=1}^{n-1} f\left(a + k \frac{b-a}{n}\right) \right]$$

denote the usual trapezoidal sum for f . Show that the error term for using T_n to estimate $\int_a^b f(x) dx$ is approximately

$$-\frac{(b-a)^2}{12n^2} [f'(b) - f'(a)].$$

Answer ☐

Exercise 429 The integral

$$\int_0^1 e^{x^2} dx = 1.462651746$$

is correct to nine decimal places. The trapezoidal rule, for $n = 1, 2$ would give

$$\int_0^1 e^{x^2} dx \approx \frac{e^0 + e^1}{2} = 1.859140914$$

and

$$\int_0^1 e^{x^2} dx \approx \frac{e^0 + 2e^{1/2} + e^1}{4} = 1.753931093.$$

At what stage in the trapezoidal rule would the approximation be correct to nine decimal places?

Answer ☐

3.10.1 Maple methods

With the advent of computer algebra packages like Maple and Mathematica one does not need to gain any expertise in computation to perform definite and indefinite integration. The reason, then, why we still drill our students on these methods is to produce an intelligent and informed user of mathematics. To illustrate here is a short Maple session on a unix computer named dogwood. After giving the maple command we are in Maple and have asked it to do some calculus questions for us. Specifically we are seeking

$$\int x^2 dx, \quad \int_0^2 x^2/dx, \quad \int \sin(4x) dx, \quad \text{and} \quad \int x[3x^2 + 2]^{5/3} dx.$$

All of these can be determined by hand using the standard methods taught for generations in calculus courses. Note that Maple is indifferent to our requirement that constants of integration should always be specified or that the interval of indefinite integration should be acknowledged.

```
[31]dogwood% maple
|^^|      Maple 12 (SUN SPARC SOLARIS)
._|_|    _/|_ . Copyright (c) Maplesoft, a division of Waterloo Maple Inc. 2008
 \ MAPLE / All rights reserved. Maple is a trademark of
 <__ __>   Waterloo Maple Inc.
 |         Type ? for help.
> int(x^2,x);
```

$$\frac{x^3}{3}$$

```
> int(x^2,x=0..2);
```

8/3

```
> int(sin(4*x),x);
```

-1/4 cos(4 x)

```
> int(x*(3*x^2+2)^(5/3),x);
```

$$\frac{(3x^2 + 2)^{8/3}}{16}$$

If we go on to ask problems that would not normally be asked on a calculus examination then the answer may be more surprising. There is no simple expression of the indefinite integral $\int \cos x^3 dx$ and consequently Maple will not find a method. The first try to obtain a precise value for $\int_0^1 \cos x^3 dx$ produces

```
> int(cos(x^3),x=0..1);
```

memory used=3.8MB, alloc=3.0MB, time=0.36

memory used=7.6MB, alloc=5.4MB, time=0.77

$$\begin{aligned} & \frac{1}{6} \pi^{1/2} \frac{(1/3)^{1/2}}{2} \left| \frac{\sin(1)^{2/3}}{30/7} - \frac{\sin(1)^{2/3} (-3/2 \cos(1) + 3/2 \sin(1))}{\pi^{1/2}} \right. \\ & \left. - \frac{\sin(1)^{2/3} \text{LommelS1}(11/6, 3/2, 1)}{\pi^{1/2}} - \frac{\sin(1)^{2/3} (\cos(1) - \sin(1)) \text{LommelS1}(5/6, 1/2, 1)}{\pi^{1/2}} \right| \end{aligned}$$

The second try asks Maple to give a numerical approximation. Maple uses a numerical integration routine with automatic error control to evaluate definite integrals that it cannot do analytically.

```
> evalf(int(cos(x^3),x=0..1));
```

0.9317044407

Thus we can be assured that $\int_0^1 \cos x^3 dx = 0.9317044407$ correct to 10 decimal places.

In short, with access to such computer methods, we can be sure that our time in studying integration theory is best spent on learning the theory so that we will understand what we are doing when we ask a computer to make calculations for us.

3.10.2 Maple and infinite integrals

For numerical computations of infinite integrals one can again turn to computer algebra packages. Here is a short Maple session that computes the infinite integrals

$$\int_0^\infty e^{-x} dx, \int_0^\infty x e^{-x} dx, \int_0^\infty x^3 e^{-x} dx, \text{ and } \int_0^\infty x^{10} e^{-x} dx.$$

We have all the tools to do these by hand, but computer methods are rather faster.

```
[32]dogwood% maple
      |\^/|      Maple 12 (SUN SPARC SOLARIS)
._|\|  |/_|. Copyright (c) Maplesoft, a division of Waterloo Maple Inc. 2008
 \  MAPLE / All rights reserved. Maple is a trademark of
<____ ____> Waterloo Maple Inc.
      |      Type ? for help.
> int( exp(-x), x=0..infinity );

1

> int(x* exp(-x), x=0..infinity );

1

> int(x^3* exp(-x), x=0..infinity );

6

> int(x^10* exp(-x), x=0..infinity );

3628800
```

Exercise 430 Show that $\int_0^\infty x^n e^{-x} dx = n!$.

Answer ☐

3.11 More Exercises

Exercise 431 If f is continuous on an interval $[a, b]$ and

$$\int_a^b f(x)g(x) dx = 0$$

for every continuous function g on $[a, b]$ show that f is identically equal to zero there. ☐

Exercise 432 ((Cauchy-Schwarz inequality)) If f and g are continuous on an interval $[a, b]$ show that

$$\left(\int_a^b f(x)g(x) dx \right)^2 \leq \left(\int_a^b [f(x)]^2 dx \right) \left(\int_a^b [g(x)]^2 dx \right).$$

Answer ☐

Exercise 433 In elementary calculus classes it is sometimes convenient to define the natural logarithm by using the integration theory,

$$\log x = \int_1^x \frac{1}{t} dt.$$

Taking this as a definition, not a computation, use the properties of integrals to develop the properties of the logarithm function. ☐

Exercise 434 Let f be a continuous function on $[1, \infty)$ such that $\lim_{x \rightarrow \infty} f(x) = \alpha$. Show that if the integral $\int_1^\infty f(x) dx$ converges, then α must be 0. ☐

Exercise 435 Let f be a continuous function on $[1, \infty)$ such that the integral $\int_1^\infty f(x) dx$ converges. Can you conclude that $\lim_{x \rightarrow \infty} f(x) = 0$? ☐

Chapter 4

Beyond the calculus integral

Our goal in this chapter is to develop the modern integral by allowing more functions to be integrated. We still insist on the viewpoint that

$$\int_a^b F'(x) dx = F(b) - F(a)$$

but we wish to relax our assumptions to allow this formula to hold even when there are infinitely many points of nondifferentiability of F . There may seem not too much point in allowing more functions to be integrated, except perhaps when one encounters a function without an integral where one seems to be needed.

But, in fact, the theory itself demands it. Too many processes of analysis lead from integrable functions [in the calculus sense] to functions for which a broader theory of integration is required. The modern theory is an indispensable tool of analysis and its theory is elegant and complete.

Remember that, for the calculus integral, an integrable function f must have an indefinite integral F for which $F'(x) = f(x)$ at every point of an interval *with finitely many exceptions*. The path to generalization is to allow infinitely many exceptional points where the derivative $F'(x)$ may not exist or may not agree with $f(x)$.

Although we will allow an infinite set, we cannot allow too large a set of exceptions. In addition, as we will find, we must impose some restrictions on the function F if we do allow an infinite set of exceptions. Those two ideas will drive the theory.

4.1 Countable sets

The first notion, historically, of a concept that captures the smallness of an infinite set is due to Cantor.

Definition 4.1 *A set of real numbers is countable if there is a sequence of real numbers $r_1, r_2, r_3, \dots$ that contains every element of the set.*

Exercise 436 *Prove that the empty set is countable.*

Answer ☐

Exercise 437 *Prove that every finite set is countable.*

Answer ☐

Exercise 438 *Prove that every subset of a countable set is countable.*

Answer ☐

Exercise 439 *Prove that the set of all integers (positive, negative or zero) is countable.*

Answer ☐

Exercise 440 *Prove that the set of all rational numbers is countable.*

Answer ☐

Exercise 441 *Prove that the union of two countable sets is countable.*

Answer ☐

Exercise 442 *Prove that the union of a sequence of countable sets is countable.*

Answer ☐

Exercise 443 *Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is a monotonic, nondecreasing function. Show that such a function may have many points of discontinuity but that the collection of all points where F is not continuous is countable.*

Answer ☐

Exercise 444 *If a function $F : (a, b) \rightarrow \mathbb{R}$ has a right-hand derivative and a left-hand derivative at a point x_0 and the derivatives on the two sides are different, then that point is said to be a corner. Show that a function may have many corners but that the collection of all corners is countable.*

☐

4.1.1 Cantor's theorem

Your first impression might be that few sets would be able to be the range of a sequence. But having seen in Exercise 440 that even the set of rational numbers that is seemingly so large can be listed, it might then appear that all sets can be so listed. After all, can you conceive of a set that is “larger” than the rationals in some way that would stop it being listed? The remarkable fact that there are sets that cannot be arranged to form the elements of some sequence was proved by Georg Cantor (1845–1918).

Theorem 4.2 (Cantor) *No interval of real numbers is countable.*

The proof is given in the next few exercises.

Exercise 445 *Prove that there would exist a countable interval if and only if the open interval $(0, 1)$ is itself countable.*

Answer ☐

Exercise 446 *Prove that the open interval $(0, 1)$ is not countable, using (as Cantor himself did) properties of infinite decimal expansions to construct a proof.*

Answer ☐

Exercise 447 *Some novices, on reading the proof of Cantor's theorem, say “Why can't you just put the number c that you found at the front of the list.” What is your rejoinder?*

Answer ☐

Exercise 448 *Give a proof that the interval (a, b) is not countable using the nested sequence of intervals argument.*

Answer ☐

Exercise 449 *We define a real number to be algebraic if it is a solution of some polynomial equation*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0,$$

where all the coefficients are integers. Thus $\sqrt{2}$ is algebraic because it is a solution of $x^2 - 2 = 0$. The number π is not algebraic because no such polynomial equation can ever be found (although this is hard to prove). Show that the set of algebraic numbers is countable.

Answer ☐

Exercise 450 *A real number that is not algebraic is said to be transcendental. For example, it is known that e and π are transcendental. What can you say about the existence of other transcendental numbers?*

Answer ☐

4.2 Derivatives which vanish outside of countable sets

Our first attempt to extend the indefinite and definite integral to handle a broader class of functions is to introduce a countable exceptional set into the definitions. We have used finite exceptional sets up to this point. This will produce a more general integral.

The principle is the following: if F is a continuous function on an interval I and if $F'(x) = 0$ for all but countably many points in I then F must be constant. The proof has appeared in Section 1.10.5.

Theorem 4.3 *Let $F : (a, b) \rightarrow \mathbb{R}$ be a function that is continuous at every point in an open interval (a, b) and suppose that $F'(x) = 0$ for all $x \in (a, b)$ with possibly countably many exceptions. Then F is a constant function.*

4.2.1 Calculus integral [countable set version]

Our original calculus integral was defined in way that was entirely dependent on the simple fact that continuous functions that have a zero derivative at all but a finite number of points must be constant. We now know that that continuous functions that have a zero derivative at all but a countable number of points must also be constant. Thus there is no reason not to extend the calculus integral to allow a countable exceptional set.

Definition 4.4 *The following describes an extension of our integration theory:*

- f is a function defined at each point of a bounded open interval (a, b) with possibly countably many exceptions.
- f is the derivative of some function in this sense: there exists a uniformly continuous function $F : (a, b) \rightarrow \mathbb{R}$ with the property that $F'(x) = f(x)$ for all $a < x < b$ with at most a countable number of exceptions.
- Then the function f is said to be integrable [in the new sense] and the value of the integral is determined by

$$\int_a^b f(x) dx = F(b-) - F(a+).$$

Zakon's Analysis text. There is currently at least one analysis textbook available¹ that follows exactly this program, replacing the Riemann integral by the Newton integral (with countably many exceptions):

Mathematical Analysis I, by Elias Zakon, ISBN 1-931705-02-X, published by The Trillia Group, 2004. 355+xii pages, 554 exercises, 26 figures, hypertextual cross-references, hyperlinked index of terms. Download size: 2088 to 2298 KB, depending on format.

This can be downloaded freely from the web site

www.trillia.com/zakon-analysisI.html

Inexpensive site licences are available for instructors wishing to adopt the text.

Zakon's text offers a serious analysis course at the pre-measure theory level, and commits itself to the Newton integral. There are rigorous proofs and the presentation is carried far enough to establish that all regulated² functions are integrable in this sense.

Exercise 451 Show that the countable set version of the calculus integral determines a unique value for the integral, i.e., does not depend on the particular antiderivative F chosen. Answer $\square$

Exercise 452 In Exercise 217 we asked the following:

Define a function $F : [0, 1] \rightarrow \mathbb{R}$ in such a way that $F(0) = 0$, and for each odd integer $n = 1, 3, 5, \dots$, $F(1/n) = 1/n$ and each even integer $n = 2, 4, 6, \dots$, $F(1/n) = 0$. On the intervals $[1/(n+1), 1/n]$ for $n = 1, 2, 3$, the function is linear. Show that $\int_a^b F'(x) dx$ exists as a calculus integral for all $0 < a < b \leq 1$ but that $\int_0^1 F'(x) dx$ does not.

Show that the new version of the calculus integral would handle this easily. Answer $\square$

¹I am indebted to Bradley Lucier, the founder of the Trillia Group, for this reference. The text has been used by him successfully for beginning graduate students at Purdue.

²Since regulated functions are uniform limits of step functions, it is easy to anticipate how this could be done. Also, regulated functions are bounded and have only a countable set of points of discontinuity. So our methods would establish integrability in this sense. One could easily rewrite this text to use the Zakon integral instead of the calculus integral. We won't.

Exercise 453 Show that all bounded functions with a countable number of discontinuities must be integrable in this new sense. Answer $\square$

Exercise 454 Show that the new version of the integral has a property that the finite set version of the integral did not have: if f_n is a sequence of functions converging uniformly to a function f on $[a, b]$ and if each f_n is integrable on $[a, b]$ then the function f must be integrable there too. $\square$

Exercise 455 Rewrite the text to use the countable set version of the integral rather than the more restrictive finite set version. Answer $\square$

Exercise 456 (limitations of the calculus integrals) Find an example of a sequence of nonnegative, integrable functions $g_k(x)$ on the interval $[0, 1]$ such that

$$\sum_{k=1}^{\infty} \left(\int_a^b g_k(x) dx \right)$$

is convergent, and yet $f = \sum_{k=1}^{\infty} g_k$ is not integrable in the calculus sense for either the finite set or countable set version.

[Note: this function *should* be integrable and the value of this integral *should* be the sum of the series. The only difficulty is that we cannot integrate enough functions. The Riemann integral has the same defect; the integral introduced later on does not. Answer $\square$]

4.3 Sets of measure zero

We shall go beyond countable sets in our search for a suitable class of small sets. A set is countable if it is small in the sense of “counting.” This is because we have defined a set to be countable if we can list off the elements of the set in the same way we list off all the *counting* numbers (i.e., 1, 2, 3, 4, ...).

We introduce a larger class of sets that is small in the sense of measuring; here we mean measuring the same way that we measure the length of an interval $[a, b]$ by the number $b - a$.

Our sets of measure zero are defined using subpartitions and very simple Riemann sums. Later on in Part Two we will find several characterizations of this important class of sets.

4.3. SETS OF MEASURE ZERO

153

Definition 4.5 A set N is said to be a set of measure zero if for every $\varepsilon > 0$ and every point $\xi \in N$ there is a $\delta(\xi) > 0$ with the following property: whenever a subpartition

$$\{([c_i, d_i], \xi_i) : i = 1, 2, \dots, n\}$$

is given with each $\xi_i \in N$ and so that

$$0 < d_i - c_i < \delta(\xi_i) \quad (i = 1, 2, \dots, n)$$

then

$$\sum_{i=1}^n (d_i - c_i) < \varepsilon.$$

Recall that in order for the subset

$$\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$$

to be a subpartition, we require merely that the intervals $\{[a_i, b_i]\}$ do not overlap. The collection here is not necessarily a partition. Our choice of language, calling it a *subpartition*, indicates that it could be (but won't be) expanded to be a partition.

Exercise 457 Show that every finite set has measure zero.

Answer ☐

Exercise 458 Show that every countable set has measure zero.

Answer ☐

Exercise 459 Show that no interval has measure zero.

Answer ☐

Exercise 460 Show that every subset of a set of measure zero must have measure zero.

Answer ☐

Exercise 461 Show that the union of two sets of measure zero must have measure zero.

Answer ☐

Exercise 462 Show that the union of a sequence of sets of measure zero must have measure zero.

Answer ☐

Exercise 463 Suppose that $\{(a_k, b_k)\}$ is a sequence of open intervals and that

$$\sum_{k=1}^{\infty} (b_k - a_k) < \infty.$$

If E is a set and every point in E belongs to infinitely many of the intervals $\{(a_k, b_k)\}$, show that E must have measure zero. Answer $\square$

4.3.1 The Cantor dust

In order to appreciate exactly what we intend by a set of measure zero we shall introduce a classically important example of such a set: the Cantor ternary set. Mathematicians who are fond of the fractal language call this set the *Cantor dust*. This suggestive phrase captures the fact that the Cantor set is indeed truly small even though it is large in the sense of counting; it is measure zero but uncountable.

We begin with the closed interval $[0, 1]$. From this interval we shall remove a dense open set G . It is easiest to understand the set G if we construct it in stages. Let $G_1 = (\frac{1}{3}, \frac{2}{3})$, and let $K_1 = [0, 1] \setminus G_1$. Thus

$$K_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$$

is what remains when the middle third of the interval $[0, 1]$ is removed. This is the first stage of our construction.

We repeat this construction on each of the two component intervals of K_1 . Let $G_2 = (\frac{1}{9}, \frac{2}{9}) \cup (\frac{7}{9}, \frac{8}{9})$ and let $K_2 = [0, 1] \setminus (G_1 \cup G_2)$. Thus

$$K_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

This completes the second stage.

We continue inductively, obtaining two sequences of sets, $\{K_n\}$ and $\{G_n\}$. The set K obtained by removing from $[0, 1]$ all of the open sets G_n is called the *Cantor set*. Because of its construction, it is often called the Cantor middle third set. In an exercise we shall present a purely arithmetic description of the Cantor set that suggests another common name for K , the *Cantor ternary set*. Figure 6.1 shows K_1 , K_2 , and K_3 .

We might mention here that variations in the constructions of K can lead to interesting situations. For example, by

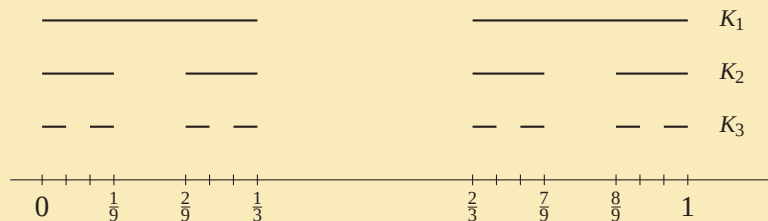


Figure 4.1: The third stage in the construction of the Cantor ternary set.

changing the construction slightly, we can remove intervals in such a way that

$$G' = \bigcup_{k=1}^{\infty} (a'_k, b'_k)$$

with

$$\sum_{k=1}^{\infty} (b'_k - a'_k) = 1/2$$

(instead of 1), while still keeping $K' = [0, 1] \setminus G'$ nowhere dense and perfect. The resulting set K' created problems for late nineteenth-century mathematicians trying to develop a theory of measure. The “measure” of G' should be $1/2$; the “measure” of $[0, 1]$ should be 1. Intuition requires that the measure of the nowhere dense set K' should be $1 - \frac{1}{2} = \frac{1}{2}$. How can this be when K' is so “small?”

Exercise 464 We have given explicit statements for K_1 and K_2 ,

$$K_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$$

and

$$K_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

What is K_3 ?

Answer $\square$

Exercise 465 Show that if this process is continued inductively, we obtain two sequences of sets, $\{K_n\}$ and $\{G_n\}$ with the following properties: For each natural number n

1. G_n is a union of 2^{n-1} pairwise disjoint open intervals.
2. K_n is a union of 2^n pairwise disjoint closed intervals.
3. $K_n = [0, 1] \setminus (G_1 \cup G_2 \cup \cdots \cup G_n)$.
4. Each component of G_{n+1} is the “middle third” of some component of K_n .
5. The length of each component of K_n is $1/3^n$.

□

Exercise 466 Establish the following observations:

1. G is an open dense set in $[0, 1]$.
2. Describe the intervals complementary to the Cantor set.
3. Describe the endpoints of the complementary intervals.
4. Show that the remaining set $K = [0, 1] \setminus G$ is closed and nowhere dense in $[0, 1]$.
5. Show that K has no isolated points and is nonempty.
6. Show that K is a nonempty, nowhere dense perfect subset of $[0, 1]$.

Answer □

Exercise 467 Show that each component interval of the set G_n has length $1/3^n$. Using this, determine that the sum of the lengths of all component intervals of G , the set removed from $[0, 1]$, is 1. Thus it appears that all of the length inside the interval $[0, 1]$ has been removed leaving “nothing” remaining.

Answer □

Exercise 468 Show that the Cantor set is a set of measure zero.

Answer □

4.3. SETS OF MEASURE ZERO

157

Exercise 469 Let E be the set of endpoints of intervals complementary to the Cantor set K . Prove that the closure of the set E is the set K . □

Exercise 470 Let G be a dense open subset of real numbers and let $\{(a_k, b_k)\}$ be its set of component intervals. Prove that $H = \mathbb{R} \setminus G$ is perfect if and only if no two of these intervals have common endpoints. □

Exercise 471 Let K be the Cantor set and let $\{(a_k, b_k)\}$ be the sequence of intervals complementary to K in $[0, 1]$. For each integer k let $c_k = (a_k + b_k)/2$ (the midpoint of the interval (a_k, b_k)) and let N be the set of points c_k for integers k . Prove each of the following:

1. Every point of N is isolated.
 2. If $c_i \neq c_j$, there exists an integer k such that c_k is between c_i and c_j (i.e., no point in N has an immediate “neighbor” in N).
-

Exercise 472 Show that the Cantor dust K can be described arithmetically as the set

$$\{x = .a_1a_2a_3 \dots \text{ (base three)} : a_i = 0 \text{ or } 2 \text{ for each } i = 1, 2, 3, \dots\}.$$

Answer □

Exercise 473 Show that the Cantor dust is an uncountable set.

Answer □

Exercise 474 Find a specific irrational number in the Cantor ternary set.

Answer □

Exercise 475 Show that the Cantor ternary set can be defined as

$$K = \left\{ x \in [0, 1] : x = \sum_{n=1}^{\infty} \frac{i_n}{3^n} \text{ for } i_n = 0 \text{ or } 2 \right\}.$$

□

Exercise 476 Let

$$D = \left\{ x \in [0, 1] : x = \sum_{n=1}^{\infty} \frac{j_n}{3^n} \text{ for } j_n = 0 \text{ or } 1 \right\}.$$

Show that $D + D = \{x + y : x, y \in D\} = [0, 1]$. From this deduce, for the Cantor ternary set K , that $K + K = [0, 2]$. □

Exercise 477 A careless student makes the following argument. Explain the error.

“If $G = (a, b)$, then $\overline{G} = [a, b]$. Similarly, if $G = \bigcup_{i=1}^{\infty} (a_i, b_i)$ is an open set, then $\overline{G} = \bigcup_{i=1}^{\infty} [a_i, b_i]$. It follows that an open set G and its closure $\overline{G}$ differ by at most a countable set. The closure just adds in all the endpoints.”

Answer □

4.4 The Devil’s staircase

The Cantor set allows the construction of a rather bizarre function that is continuous and nondecreasing on the interval $[0, 1]$. It has the property that it is constant on every interval complementary to the Cantor set and yet manages to increase from $f(0) = 0$ to $f(1) = 1$ by doing all of its increasing on the Cantor set itself. It has sometimes been called “the devil’s staircase” or simply the Cantor function.

Thus this is an example of a continuous function on the interval $[0, 1]$ which has a zero derivative everywhere outside of the Cantor set. If we were to try to develop a theory of indefinite integration that allows exceptional sets of measure zero we would have to impose some condition that excludes such functions. We will see that condition in Section 4.5.4.

4.4.1 Construction of Cantor’s function

Define the function f in the following way. On the open interval $(\frac{1}{3}, \frac{2}{3})$, let $f = \frac{1}{2}$; on the interval $(\frac{1}{9}, \frac{2}{9})$, let $f = \frac{1}{4}$; on $(\frac{7}{9}, \frac{8}{9})$, let $f = \frac{3}{4}$. Proceed inductively. On the 2^{n-1} open intervals appearing at the n th stage of our construction of the Cantor set, define f to satisfy the following conditions:

1. f is constant on each of these intervals.

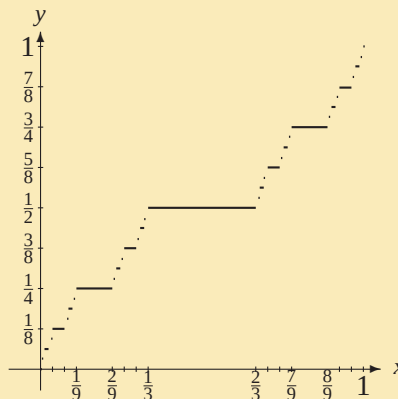


Figure 4.2: The third stage in the construction of the Cantor function.

2. f takes the values

$$\frac{1}{2^n}, \frac{3}{2^n}, \dots, \frac{2^n - 1}{2^n}$$

on these intervals.

3. If x and y are members of different n th-stage intervals with $x < y$, then $f(x) < f(y)$.

This description defines f on $G = [0, 1] \setminus K$. Extend f to all of $[0, 1]$ by defining $f(0) = 0$ and, for $0 < x \leq 1$,

$$f(x) = \sup\{f(t) : t \in G, t < x\}.$$

Figure 4.2 illustrates the initial stages of the construction. The function f is called the *Cantor function*. Observe that f “does all its rising” on the set K .

The Cantor function allows a negative answer to many questions that might be asked about functions and derivatives and, hence, has become a popular counterexample. For example, let us follow this kind of reasoning. If f is a continuous function on $[0, 1]$ and $f'(x) = 0$ for every $x \in (0, 1)$ then f is constant. (This is proved in most calculus courses by using the mean value theorem.) Now suppose that we know less, that $f'(x) = 0$ for every $x \in (0, 1)$ excepting a “small” set E

of points at which we know nothing. If E is finite it is still easy to show that f must be constant. If E is countable it is possible, but a bit more difficult, to show that it is still true that f must be constant. The question then arises, just how small a set E can appear here; that is, what would we have to know about a set E so that we could say $f'(x) = 0$ for every $x \in (0, 1) \setminus E$ implies that f is constant?

The Cantor function is an example of a function constant on every interval complementary to the Cantor set K (and so with a zero derivative at those points) and yet is not constant. The Cantor set, since it is both measure zero and nowhere dense, might be viewed as extremely small, but even so it is not insignificant for this problem.

Exercise 478 *In the construction of the Cantor function complete the verification of details.*

1. Show that $f(G)$ is dense in $[0, 1]$.
2. Show that f is nondecreasing on $[0, 1]$.
3. Infer from (a) and (b) that f is continuous on $[0, 1]$.
4. Show that $f(K) = [0, 1]$ and thus (again) conclude that K is uncountable.

□

Exercise 479 *Show that the Cantor function has a zero derivative everywhere on the open set complementary to the Cantor set in the interval $[0, 1]$. [In more colorful language, we say that this function has a zero derivative almost everywhere.]*

□

Exercise 480 *Each number x in the Cantor set can be written in the form*

$$x = \sum_{i=1}^{\infty} 23^{-n_i}$$

for some increasing sequence of integers $n_1 < n_2 < n_3 < \dots$. Show that the Cantor function assumes the value $F(x) = \sum_{i=1}^{\infty} 2^{-n_i}$ at each such point.

□

Exercise 481 *Show that the Cantor function is a monotone, nondecreasing function on $[0, 1]$ that has these properties:*

1. $F(0) = 0$,

$$2. F(x/3) = F(x)/2,$$

$$3. F(1-x) = 1 - F(x).$$

[In fact the Cantor function is the only monotone, nondecreasing function on $[0, 1]$ that has these three properties.]

Answer $\square$

4.5 Functions with zero variation

Sets of measure zero have been defined by requiring certain small sums

$$\sum_{i=1}^n (b_i - a_i)$$

whenever a subpartition

$$\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$$

is controlled by a function $\delta(x)$. We are interested in other variants on this same theme, involving sums of the form

$$\sum_{i=1}^n |F(b_i) - F(a_i)| \text{ or } \sum_{i=1}^n |f(\xi_i)|(b_i - a_i) \text{ or even } \sum_{i=1}^n |F(b_i) - F(a_i) - f(\xi_i)|(b_i - a_i)|.$$

A measurement of the sums

$$\sum_{i=1}^n |F(b_i) - F(a_i)|$$

taken over nonoverlapping subintervals is considered to compute the *variation* of the function F . This notion appears in the early literature and was formalized by Jordan in the late 19th century under the terminology “variation of a function.”

We do not need the actual measurement of variation. What we do need is the notion that a function has *zero variation*. This is a function that has only a small change on a set, or whose growth on the set is insubstantial.

Definition 4.6 A function $F : (a, b) \rightarrow \mathbb{R}$ is said to have zero variation on a set $E \subset (a, b)$ if for every $\varepsilon > 0$ and every $x \in E$ there is a $\delta(x) > 0$

$$\sum_{i=1}^n |F(b_i) - F(a_i)| < \varepsilon$$

whenever a subpartition $\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$ is chosen for which

$$\xi_i \in E \cap [a_i, b_i] \text{ and } b_i - a_i < \delta(\xi_i).$$

We saw a definition very similar to this when we defined a set of measure zero. In fact the formal nature of the definition is exactly the same as the requirement that a set E should have measure zero. The following exercise makes this explicit. As we shall discover, all of the familiar functions of the calculus turn out to have zero variation on sets of measure zero. Only rather pathological examples (notably the Cantor function) do not have this property.

Exercise 482 Show that a set E has measure zero if and only if the function $F(x) = x$ has zero variation on E .

Answer $\square$

Exercise 483 Suppose that $F : \mathbb{R} \rightarrow \mathbb{R}$ has zero variation on a set E_1 and that $E_2 \subset E_1$. Show that then F has zero variation on E_2 .

Answer $\square$

Exercise 484 Suppose that $F : \mathbb{R} \rightarrow \mathbb{R}$ has zero variation on the sets E_1 and E_2 . Show that then F has zero variation on the union $E_1 \cup E_2$.

Answer $\square$

Exercise 485 Suppose that $F : \mathbb{R} \rightarrow \mathbb{R}$ has zero variation on each member of a sequence of sets $E_1, E_2, E_3, \dots$. Show that then F has zero variation on the union $\bigcup_{n=1}^{\infty} E_n$.

Answer $\square$

Exercise 486 Prove the following theorem that shows another important version of zero variation. We could also describe this as showing a function has small Riemann sums over sets of measure zero.

Theorem 4.7 Let f be defined at every point of a measure zero set N and let $\varepsilon > 0$. Then for every $x \in N$ there is a $\delta(x) > 0$ so that

$$\sum_{i=1}^n |f(\xi_i)|(b_i - a_i) < \varepsilon$$

whenever a subpartition $\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$ is chosen for which $\xi_i \in N \cap [a_i, b_i]$ and $b_i - a_i < \delta(\xi_i)$.

Answer $\square$

Exercise 487 Let F be defined on an open interval (a, b) and let f be defined at every point of a measure zero set $N \subset (a, b)$. Suppose that F has zero variation on N . Let $\varepsilon > 0$. Show for every $x \in N$ there is a $\delta(x) > 0$ such that

$$\sum_{i=1}^n |F(b_i) - F(a_i) - f(\xi_i)(b_i - a_i)| < \varepsilon$$

whenever a subpartition $\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$ is chosen for which

Answer $\square$

Exercise 488 Let F be defined on an open interval (a, b) and let f be defined at every point of a set E . Suppose that $F'(x) = f(x)$ for every $x \in E$. Let $\varepsilon > 0$. Show for every $x \in E$ there is a $\delta(x) > 0$ such that

$$\sum_{i=1}^n |F(b_i) - F(a_i) - f(\xi_i)(b_i - a_i)| < \varepsilon$$

whenever a subpartition $\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$ is chosen for which

$$\xi_i \in E \cap [a_i, b_i] \text{ and } b_i - a_i < \delta(\xi_i).$$

Answer $\square$

Exercise 489 Show that the Cantor function has zero variation on the open set complementary to the Cantor set in the interval $[0, 1]$.

Answer $\square$

4.5.1 Zero variation lemma

The fundamental growth theorem that we need shows that only constant functions have zero variation on an interval.

Theorem 4.8 Suppose that a function $F : (a, b) \rightarrow \mathbb{R}$ has zero variation on the entire interval (a, b) . Then F is constant on that interval.

Exercise 490 Use a Cousin covering argument to prove the theorem.

Answer ☐

Exercise 491 Show that the Cantor function does not have zero variation on the Cantor set.

Answer ☐

4.5.2 Zero derivatives imply zero variation

There is an immediate connection between the derivative and its variation in a set. In the simplest case we see that a function has zero variation on a set on which it has everywhere a zero derivative.

Theorem 4.9 Suppose that a function $F : (a, b) \rightarrow \mathbb{R}$ has a zero derivative $F'(x)$ at every point x of a set $E \subset (a, b)$. Then F has zero variation on E .

Exercise 492 Prove Theorem 4.9 by applying Exercise 488.

☐

Exercise 493 Give a direct proof of Theorem 4.9.

Answer ☐

4.5.3 Continuity and zero variation

There is an intimate and immediate relation between continuity and zero variation.

Theorem 4.10 Suppose $F : (a, b) \rightarrow \mathbb{R}$. Then F is continuous at a point $x_0 \in (a, b)$ if and only if F has zero variation on the singleton set $E = \{x_0\}$.

Corollary 4.11 Suppose $F : (a, b) \rightarrow \mathbb{R}$. Then F is continuous at each point $c_1, c_2, c_3, \dots, c_k \in (a, b)$ if and only if F has zero variation on the finite set $E = \{c_1, c_2, c_3, \dots, c_k\}$.

Corollary 4.12 Suppose $F : (a, b) \rightarrow \mathbb{R}$. Then F is continuous at each point $c_1, c_2, c_3, \dots$ from a sequence of points in (a, b) if and only if F has zero variation on the countable set $E = \{c_1, c_2, c_3, \dots\}$.

Exercise 494 Suppose $F : (a, b) \rightarrow \mathbb{R}$. Show that F is continuous at every point in a set E if and only if F has zero variation in every countable subset of E . □

4.5.4 Absolute continuity

We have seen that the function $F(x) = x$ has zero variation on a set N precisely when that set N is a set of measure zero. We see, then, that F has zero variation on all sets of measure zero. Most functions that we have encountered in the calculus also have this property. We shall see that all differentiable functions have this property. It plays a vital role in the theory; such functions are said to be absolutely continuous³.

Definition 4.13 A function $F : (a, b) \rightarrow \mathbb{R}$ is said to be absolutely continuous on the open interval (a, b) if F has zero variation on every subset N of the interval that has measure zero.

The exercises show that most continuous functions we encounter in the calculus will be absolutely continuous. In fact the only continuous function we have seen so far that is not absolutely continuous is the Cantor function.

Exercise 495 Show that the function $F(x) = x$ is absolutely continuous on every open interval. □

Exercise 496 Show that a linear combination of absolutely continuous functions is absolutely continuous. □

Exercise 497 Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is absolutely continuous on the interval (a, b) . Show that F must be continuous at every point of that interval. □

Exercise 498 Show that a Lipschitz function defined on an open interval is absolutely continuous there. □

³**Note to the instructor:** this notion is strictly more general than the traditional notion (due to Vitali) of a function absolutely continuous on a closed, bounded interval $[a, b]$. In particular an absolutely continuous function in this sense need not have bounded variation. See Exercise 4.14

Exercise 499 Give an example of an absolutely continuous function that is not Lipschitz. □

Exercise 500 Show that the Cantor function is not absolutely continuous on $(0, 1)$. □

Exercise 501 Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is differentiable at each point of the open interval (a, b) . Show that F is absolutely continuous on the interval (a, b) . □

Exercise 502 Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is differentiable at each point of the open interval (a, b) with finitely many exceptions but that F is continuous at those exceptional points. Show that F is absolutely continuous on the interval (a, b) . □

Exercise 503 Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is differentiable at each point of the open interval (a, b) with countably many exceptions but that F is continuous at those exceptional points. Show that F is absolutely continuous on the interval (a, b) . □

Exercise 504 Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is differentiable at each point of the open interval (a, b) with the exception of a set $N \subset (a, b)$. Suppose further that N is a set of measure zero and that F has zero variation on N . Show that F is absolutely continuous on the interval (a, b) . □

Exercise 505 Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is absolutely continuous on the interval (a, b) . Then by definition F has zero variation on every subset of measure zero. Is it possible that F has zero variation on subsets that are not measure zero? □

Exercise 506 A function F on an open interval I is said to have finite derived numbers on a set $E \subset I$ if, for each $x \in E$, there is a number M_x and one can choose $\delta > 0$ so that

$$\left| \frac{F(x+h) - F(x)}{h} \right| \leq M_x$$

whenever $x+h \in I$ and $|h| < \delta$. Show that F is absolutely continuous on E if F has finite derived numbers there. [cf. Exercise 170.] □

4.5.5 Absolute continuity in Vitali's sense

There is a type of absolute continuity, due to Vitali, that is very similar to the ϵ - δ definition of uniform continuity.

Definition 4.14 (Absolute continuity in Vitali's sense) A function $F : [a, b] \rightarrow \mathbb{R}$ is absolutely continuous in Vitali's sense on $[a, b]$ provided that for every $\epsilon > 0$ there is a $\delta > 0$ so that

$$\sum_{i=1}^n |F(x_i) - F(y_i)| < \epsilon$$

whenever $\{[x_i, y_i]\}$ are nonoverlapping subintervals of $[a, b]$ for which

$$\sum_{i=1}^n (y_i - x_i) < \delta.$$

This condition is strictly stronger than absolute continuity: there are absolutely continuous functions that are not absolutely continuous in Vitali's sense.

Exercise 507 Prove: If F is absolutely continuous in Vitali's sense on $[a, b]$ then F is uniformly continuous there. □

Exercise 508 Prove: If F is absolutely continuous in Vitali's sense on $[a, b]$ then F is absolutely continuous on the open interval (a, b) . □

Exercise 509 Prove: If F is absolutely continuous in Vitali's sense on $[a, b]$ then F has bounded variation on $[a, b]$. □

Exercise 510 If F is Lipschitz show that F is absolutely continuous in Vitali's sense. □

Exercise 511 Show that an everywhere differentiable function must be absolutely continuous on any interval (a, b) but need not be absolutely continuous in Vitali's sense on $[a, b]$. □

4.6 The integral

Our theory so far in Part One has introduced and studied the calculus integral, both as indefinite and definite integral. The key point in that theory was simply this observation:

Continuous functions on open intervals whose derivatives are determined at all but finitely many points are unique up to an additive constant.

The whole theory of the calculus integral was based on this simple concept. We can consider that this simple phrase is enough to explain the elementary theory of integration.

The exceptional set that we allowed was always finite. To go beyond that and provide a more comprehensive integration theory we must allow infinite sets. We have seen that sets of measure zero offer a useful class of exceptional sets. But we also saw the Cantor function whose derivative is zero everywhere except on the measure zero Cantor set, and yet the Cantor function is not constant.

Absolutely continuous functions behave on sets of measure zero in precisely the manner that we require. To avoid pathological functions like the Cantor function we need to assume some kind of absolute continuity or we need to assume that the functions we use have zero variation on certain sets. Thus we can build a new theory of integration on a statement that generalizes the one above:

Absolutely continuous functions on open intervals whose derivatives are determined at all but a set of measure zero are unique up to an additive constant.

We can consider that this simple phrase, too, is enough to explain the modern theory of integration. We formulate our definitions in a way that mimics and extends the calculus integral, taking advantage now of sets of measure zero and absolutely continuous functions.

Definition 4.15 (Definite integral) Let $f : (a, b) \rightarrow \mathbb{R}$ be a function defined at all points of the open interval (a, b) with the possible exception of a set of measure zero. Then f is said to be integrable on the closed, bounded interval $[a, b]$ provided there is a function $F : (a, b) \rightarrow \mathbb{R}$ so that

1. F is uniformly continuous on (a, b) .
2. F is absolutely continuous on (a, b) .
3. $F'(x) = f(x)$ at all points x of (a, b) with the possible exception of a set of measure zero.

In that case we define $\int_a^b f(x) dx = F(b-) - F(a+)$.

We recall that, because F is uniformly continuous on (a, b) , the two one-sided limits $F(b-)$ and $F(a+)$ must exist. Most often we have determined F on all of $[a, b]$ and so can simply use $F(b)$ and $F(a)$. Sometimes it is more convenient to state the conditions for the integral with direct attention to the set of exceptional points where the derivative $F'(x) = f(x)$ may fail.

Definition 4.16 (Definite integral) Let $f : (a, b) \rightarrow \mathbb{R}$ be a function defined at all points of the open interval (a, b) with the possible exception of a set of measure zero. Then f is said to be integrable on the closed, bounded interval $[a, b]$ provided there is a function $F : [a, b] \rightarrow \mathbb{R}$ and there is a set $N \subset (a, b)$ so that

1. F is uniformly continuous on (a, b) .
2. N has measure zero.
3. $F'(x) = f(x)$ at all points x of (a, b) with the possible exception of points in N .
4. F has zero variation on N .

In that case we define $\int_a^b f(x) dx = F(b-) - F(a+)$.

Exercise 512 Show that Definition 4.15 and Definition 4.16 are equivalent. □

Exercise 513 Show that the following requirements are not equivalent to those in Definition 4.15 but are stronger.

1. F is absolutely continuous in Vitali's sense on $[a, b]$.
2. $F'(x) = f(x)$ at all points x of (a, b) with the possible exception of points in a set of measure zero.
3. $\int_a^b f(x) dx = F(b) - F(a)$.

[This set of stronger requirements describes Lebesgue's integral which is less general than the integral defined here.]

Answer □

Exercise 514 Under what hypotheses is

$$\int_a^b F'(x) dx = F(b) - F(a)$$

a correct statement?

Answer □

Exercise 515 Show that the new definition of definite integral (either Definition 4.15 or Definition 4.16) includes the notion of definite integral from Chapter 3. □

Exercise 516 Show that the new definition of definite integral (either Definition 4.15 or Definition 4.16) includes, as integrable, functions that would not be considered integrable in Chapter 3. □

4.6.1 Infinite integrals

Exactly the same definition for the infinite integrals

$$\int_{-\infty}^{\infty} f(x) dx, \int_a^{\infty} f(x) dx, \text{ and } \int_{-\infty}^b f(x) dx$$

can be given as for the integral over a closed bounded interval.

Definition 4.17 Let f be a function defined at every point of (∞, ∞) with the possible exception of a set of measure zero. Then f is said to be integrable on (∞, ∞) provided there is a function $F : (-\infty, \infty) \rightarrow \mathbb{R}$ so that

1. F is absolutely continuous on (∞, ∞) .
2. $F'(x) = f(x)$ at all points x with the possible exception of a set of measure zero.
3. Both limits $F(\infty) = \lim_{x \rightarrow \infty} F(x)$ and $F(-\infty) = \lim_{x \rightarrow -\infty} F(x)$ exist.

In that case the number $\int_{-\infty}^{\infty} f(x) dx = F(\infty) - F(-\infty)$, is called the definite integral of f on the interval (∞, ∞) .

Similar assertions define

$$\int_{-\infty}^b f(x) dx = F(b) - F(-\infty)$$

and

$$\int_a^{\infty} f(x) dx = F(\infty) - F(a).$$

In analogy with the terminology of an infinite series $\sum_{k=1}^{\infty} a_k$ we often say that the integral $\int_a^{\infty} f(x) dx$ *converges* when the integral exists. That suggests language asserting that the integral *converges absolutely* if both integrals

$$\int_a^{\infty} f(x) dx \quad \text{and} \quad \int_a^{\infty} |f(x)| dx$$

exist.

4.7 Lipschitz functions and bounded integrable functions

We know that Lipschitz functions are absolutely continuous. They are even absolutely continuous in Vitali's sense which is a stronger statement. Thus the theory of integration for bounded functions can be restated in the following language.

Theorem 4.18 (Definite integral of bounded functions) *Let $f : (a, b) \rightarrow \mathbb{R}$ be a bounded function defined at all points of the open interval (a, b) with the possible exception of a set of measure zero. Then f is integrable on the closed, bounded interval $[a, b]$ if and only if there is a function $F : [a, b] \rightarrow \mathbb{R}$ so that*

1. F is Lipschitz on the closed interval $[a, b]$.
2. $F'(x) = f(x)$ at all points x of the open interval (a, b) with the possible exception of points in a set of measure zero.

In that case $\int_a^b f(x) dx = F(b) - F(a)$.

This can also be considered the definition of the classical Lebesgue integral of bounded functions. Since Lebesgue started with the bounded functions this is an historically important integral. Since a vast number of problems in integration theory consider only bounded functions this is also a reasonable working definition for any problem in integration theory that does not lead to unbounded functions.

Exercise 517 *Prove Theorem 4.18.*

□

4.8 Approximation by Riemann sums

We have seen that all calculus integrals can be approximated by Riemann sums, *pointwise approximated* that is. The same theorem is true for the advanced integration theory. In Theorem 4.19 below we see that the property of being an integral (which is a property expressed in the language of derivatives, zero measure sets and zero variation) can be completely described by a property expressed by partitions and Riemann sums.

This theorem was first observed by the Irish mathematician Ralph Henstock. Since then it has become the basis for a definition of the modern integral. The proof is elementary. Even so, it is remarkable and was not discovered until the 1950s, in spite of intense research into integration theory in the preceding half-century.

Theorem 4.19 (Henstock's criterion) *Suppose that f is an integrable function defined at every point of a closed, bounded interval $[a, b]$. Then for every $\varepsilon > 0$ and every point $x \in [a, b]$ there is a $\delta(x) > 0$ so that*

$$\sum_{i=1}^n \left| \int_{a_i}^{b_i} f(x) dx - f(\xi_i)(b_i - a_i) \right| < \varepsilon$$

and

$$\left| \int_a^b f(x) dx - \sum_{i=1}^n f(\xi_i)(b_i - a_i) \right| < \varepsilon$$

whenever a partition of the interval $[a, b]$ $\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$ is chosen for which

$$\xi_i \in [a_i, b_i] \text{ and } b_i - a_i < \delta(\xi_i).$$

This theorem is stated in only one direction: if f is integrable then the integral has a pointwise approximation using Riemann sums. The converse direction is true too and can be used to define the integral by means of Riemann sums. Of course, one is then obliged to develop the full theory of zero measure sets, zero variation and absolute continuity in order to connect the two theories and show that they are equivalent.

The theorem provides only for a pointwise approximation by Riemann sums. It is only under rather severe conditions that it is possible to find a uniform approximation by Riemann sums. Exercises 519, 520, and 521 provide that information.

Exercise 518 Prove Theorem 4.19.

Answer ☐

Exercise 519 (Riemann criterion) Theorem 4.19 shows that the integral has a pointwise approximation using Riemann sums. Show that a function f would have a uniform approximation using Riemann sums if and only if, for any $\varepsilon > 0$, there is a partition of the interval $[a, b]$ $\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$ for which

$$\sum_{i=1}^n \omega(f, [a_i, b_i])(b_i - a_i) < \varepsilon.$$

☐

Exercise 520 (Lebesgue criterion) Theorem 4.19 shows that the integral has a pointwise approximation using Riemann sums. Show that, if f is bounded and the set of points of discontinuity form a set of measure zero, then the integral has a uniform approximation using Riemann sums.

☐

Exercise 521 (Lebesgue criterion) Theorem 4.19 shows that the integral has a pointwise approximation using Riemann sums. Show that, if the integral has a uniform approximation using Riemann sums, then f must be bounded and the set of points of discontinuity must form a set of measure zero.

☐

4.9 Properties of the integral

The basic properties of integrals are easily studied for the most part since they are natural extensions of properties we have already investigated for the calculus integral. There are some surprises and some deep properties which were either false for the calculus integral or were hidden too deep for us to find without the tools we have now developed.

We know these formulas for the narrow calculus integral and we are interested in extending them to full generality. If you are working largely with continuous functions then there is little need to know just how general these properties can be developed.

4.9.1 Inequalities

Formula for inequalities:

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx$$

if $f(x) \leq g(x)$ for all points x in (a, b) except possibly points of a set of measure zero.

We have seen this statement before for the calculus integral in Section 3.4.1 where we allowed only a finite number of exceptions for the inequality. Here is a precise statement of what we intend here by this statement: If both functions $f(x)$ and $g(x)$ have an integral on the interval $[a, b]$ and, if $f(x) \leq g(x)$ for all points x in (a, b) except possibly points of a set of measure zero, then the stated inequality must hold.

Exercise 522 Complete the details needed to prove the inequality formula.

Answer $\square$

4.9.2 Linear combinations

Formula for linear combinations:

$$\int_a^b [rf(x) + sg(x)] dx = r \int_a^b f(x) dx + s \int_a^b g(x) dx \quad (r, s \in \mathbb{R}).$$

We have seen this statement before for the calculus integral in Section 3.4.2 Here is a precise statement of what we intend now by this formula: If both functions $f(x)$ and $g(x)$ have an integral on the interval $[a, b]$ then any linear combination $rf(x) + sg(x)$ ($r, s \in \mathbb{R}$) also has an integral on the interval $[a, b]$ and, moreover, the identity must hold. The proof is an exercise in derivatives, taking proper care of the exceptional sets of measure zero. We know, as usual, that

$$\frac{d}{dx} (rF(x) + sG(x)) = rF'(x) + sG'(x)$$

at any point x at which both F and G are differentiable.

Exercise 523 Complete the details needed to prove the linear combination formula.

Answer $\square$

4.9.3 Subintervals

Formula for subintervals: If $a < c < b$ then

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$$

The intention of the formula is contained in two statements in this case:

If the function $f(x)$ has an integral on the interval $[a, b]$ then $f(x)$ must also have an integral on any closed subinterval of the interval $[a, b]$ and, moreover, the identity must hold.

and

If the function $f(x)$ has an integral on the interval $[a, c]$ and also on the interval $[c, b]$ then $f(x)$ must also have an integral on the interval $[a, b]$ and, moreover, the identity must hold.

Exercise 524 Supply the details needed to prove the subinterval formula.

Answer $\square$

4.9.4 Integration by parts

Integration by parts formula:

$$\int_a^b F(x)G'(x) dx = F(b)G(b) - F(a)G(b) - \int_a^b F'(x)G(x) dx$$

The intention of the formula is contained in the product rule for derivatives:

$$\frac{d}{dx} (F(x)G(x)) = F(x)G'(x) + F'(x)G(x)$$

which holds at any point where both functions are differentiable. One must then give strong enough hypotheses that the function $F(x)G(x)$ is an indefinite integral for the function

$$F(x)G'(x) + F'(x)G(x)$$

in the sense needed for our integral.

Exercise 525 Supply the details needed to state and prove an integration by parts formula for this integral.

Answer $\square$

4.9.5 Change of variable

The change of variable formula (i.e., integration by substitution):

$$\int_a^b f(g(t))g'(t) dt = \int_{g(a)}^{g(b)} f(x) dx.$$

The proof for the calculus integral was merely an application of the chain rule for the derivative of a composite function:

$$\frac{d}{dx} F(G(x)) = F'(G(x))G'(x).$$

Since our extended integral includes the calculus integral we still have this formula for all the old familiar cases. It is possible to extend the formula to handle much more general situations.

Exercise 526 Supply the details needed to state and prove a change of variables formula for this integral. Answer $\square$

Exercise 527 (no longer failed change of variables) In Exercise 280 we discovered that the calculus integral did not permit the change of variables, $F(x) = |x|$ and $G(x) = x^2 \sin x^{-1}$, $G(0) = 0$ in the integral

$$\int_0^1 F'(G(x))G'(x) dx = F(G(1)) - F(G(0)) = |\sin 1|.$$

Is this valid now?

Answer $\square$

4.9.6 What is the derivative of the definite integral?

What is

$$\frac{d}{dx} \int_a^x f(t) dt?$$

We know that $\int_a^x f(t) dt$ is an indefinite integral of f and so, by definition,

$$\frac{d}{dx} \int_a^x f(t) dt = f(x)$$

at all points in the interval (a, b) except possibly at the points of a set of measure zero.

We can still make the same observation that we did in Section 3.4.6 for the calculus integral:

$$\frac{d}{dx} \int_a^x f(t) dt = f(x)$$

at all points $a < x < b$ at which f is continuous. But this is quite misleading here. The function may be discontinuous everywhere, and yet the differentiation formula always holds for most points x .

4.9.7 Monotone convergence theorem

For this integral we can integrate a limit of a monotone sequence by interchanging the limit and the integral.

Theorem 4.20 (Monotone convergence theorem) Let $f_n : [a, b] \rightarrow \mathbb{R}$ ($n = 1, 2, 3, \dots$) be a nondecreasing sequence of functions, each integrable on the interval $[a, b]$. Suppose that, for all x in (a, b) except possibly a set of measure zero,

$$f(x) = \lim_{n \rightarrow \infty} f_n(x).$$

Then f is integrable on $[a, b]$ and

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx$$

provided this limit exists.

The exciting part of this statement has been underlined. Unfortunately it is more convenient for us to leave the proof of this fact to a more advanced course. Thus in the exercise you are asked to prove only a weaker version.

Exercise 528 Prove the formula without the underlined statement, i.e., assume that f is integrable and then prove the identity. Answer $\square$

Exercise 529 State and prove a version of the formula

$$\int_a^b \left(\lim_{n \rightarrow \infty} f_n(x) \right) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

using uniform convergence as your main hypothesis. $\square$

4.9.8 Summation of series theorem

For this integral we can sum series of nonnegative terms and integrate term-by-term.

Theorem 4.21 (summation of series) Suppose that $g_1, g_2, g_3, \dots$ is a sequence of nonnegative functions, each one integrable on a closed bounded interval $[a, b]$. Suppose that, for all x in (a, b) except possibly a set of measure zero,

$$f(x) = \sum_{k=1}^{\infty} g_k(x).$$

Then f is integrable on $[a, b]$ and

$$\int_a^b f(x) dx = \sum_{k=1}^{\infty} \left(\int_a^b g_k(x) dx \right) \quad (4.1)$$

provided the series converges.

The exciting part of this statement, again, has been underlined. Unfortunately it is more convenient for us to leave the proof of this fact to a more advanced course. Thus in the exercise you are asked to prove only a weaker version.

Exercise 530 Prove the formula without the underlined statement, i.e., assume that f is integrable and then prove the identity. Answer $\square$

4.9.9 Null functions

A function $f : [a, b] \rightarrow \mathbb{R}$ is said to be a null function on $[a, b]$ if it is defined at almost every point of $[a, b]$ and is zero at almost every point of $[a, b]$. Thus these functions are, for all practical purposes, just the zero function. They are particularly easy to handle in this theory for that reason.

Exercise 531 Let $f : [a, b] \rightarrow \mathbb{R}$ be a null function on $[a, b]$. Then f is integrable on $[a, b]$ and

$$\int_a^b f(x) dx = 0.$$

Answer $\square$

Exercise 532 Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is an integrable function on $[a, b]$ and that

$$\int_c^d f(x) dx = 0 \quad \text{for all } a \leq c < d \leq b.$$

Then f is a null function on $[a, b]$.

Answer ☐

Exercise 533 Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is a nonnegative, integrable function on $[a, b]$ and that

$$\int_a^b f(x) dx = 0.$$

Then f is a null function on $[a, b]$.

Answer ☐

4.10 The Henstock-Kurweil integral

Definition 4.22 (Henstock-Kurzweil integral) Suppose that f is defined at every point of a closed, bounded interval $[a, b]$. Then f is said to be Henstock-Kurzweil integrable on $[a, b]$ if there is a number I with the property that, for every $\varepsilon > 0$ and every point $x \in [a, b]$ there is a $\delta(x) > 0$ so that

$$\left| I - \sum_{i=1}^n f(\xi_i)(b_i - a_i) \right| < \varepsilon$$

whenever a partition of $[a, b]$ $\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$ is chosen for which

$$\xi_i \in [a_i, b_i] \text{ and } b_i - a_i < \delta(\xi_i).$$

The number I is set equal to $\int_a^b f(x) dx$ and the latter is called the Henstock-Kurzweil integral of f on $[a, b]$.

Here are some remarks that you should be able to prove or research.

1. The Henstock-Kurzweil integral not only includes, but is equivalent to the integral defined in this chapter.
2. There are bounded, Henstock-Kurzweil integrable functions that are not integrable [calculus sense].
3. There are unbounded, Henstock-Kurzweil integrable functions that are not integrable [calculus sense].
4. The Henstock-Kurzweil integral is a nonabsolute integral, i.e., there are integrable functions f for which $|f|$ is not integrable.

5. The Henstock-Kurzweil integral is often considered to be the correct version of integration theory on the line, but one that only specialists would care to learn. (?)

There are now a number of texts that start with Definition 4.22 and develop the theory of integration on the real line in a systematic way. Too much time, however, working with the technical details of Riemann sums may not be entirely profitable since most advanced textbooks will use measure theory exclusively. Our text

[TBB] B. S. Thomson, J. B. Bruckner, A. M. Bruckner, *Elementary Real Analysis: Dripped Version*, ClassicalRealAnalysis.com (2008).

available for free at our website contains a brief account of the calculus integral and several chapters devoted to the Henstock-Kurweil integral. After that integration theory is developed we then can give a fairly rapid and intuitive account of the measure theory that most of us are expected to know by a graduate level.

4.11 The Lebesgue integral

Lebesgue gave a number of definitions for his integral; the most famous is the constructive definition using his measure theory. He also gave a descriptive definition similar to the calculus definitions that we are using in this text. For bounded functions his definition⁴ is exactly as given below. The second definition, for unbounded functions, uses the later notion due to Vitali that we have investigated in Section 4.5.5.

⁴Here is a remark on this fact from *Functional Analysis*, by Frigyes Riesz, Béla Szökefalvi-Nagy, and Leo F. Boron: “Finally, we discuss a definition of the Lebesgue integral based on differentiation, just as the classical integral was formerly defined in many textbooks of analysis. A similar definition, if only for bounded functions, was already formulated in the first edition of Lebesgue’s *Leçons sur l’intégration*, but without being followed up: ‘A bounded function $f(x)$ is said to be summable if there exists a function $F(x)$ with bounded derived numbers [i.e., Lipschitz] such that $F(x)$ has $f(x)$ for derivative, except for a set of values of x of measure zero. The integral in (a, b) is then, by definition, $F(b) - F(a)$.’

Definition 4.23 Let f be a bounded function that is defined at every point of $[a, b]$. Then, f is said to be Lebesgue integrable on $[a, b]$ if there is a Lipschitz function $F : [a, b] \rightarrow \mathbb{R}$ such that $F'(x) = f(x)$ at every point of (a, b) with the exception of points in a set of measure zero. In that case we define

$$\int_a^b f(x) dx = F(b) - F(a)$$

and this number is called the Lebesgue integral of f on $[a, b]$.

Definition 4.24 Let f be a function, bounded or unbounded, that is defined at every point of $[a, b]$. Then, f is said to be Lebesgue integrable on $[a, b]$ if there is a function $F : [a, b] \rightarrow \mathbb{R}$ that is absolutely continuous in the sense of Vitali on $[a, b]$ such that $F'(x) = f(x)$ at every point of (a, b) with the exception of points in a set of measure zero. In that case we define

$$\int_a^b f(x) dx = F(b) - F(a)$$

and this number is called the Lebesgue integral of f on $[a, b]$.

Strictly speaking the Lebesgue integral does not quite go “beyond the calculus integral.” For bounded functions, the Lebesgue integral includes the calculus integral and integrates many important classes of functions that the calculus integral cannot manage. But for unbounded functions the relation between the calculus integral and the Lebesgue integral is more delicate: there are functions integrable in the calculus sense but which are not absolutely integrable. Any one of these functions must fail to have a Lebesgue integral.

Here are some remarks that you should be able to prove or research.

1. There are unbounded, integrable functions [calculus sense] that are not Lebesgue integrable.
2. All bounded, integrable functions [calculus sense] are Lebesgue integrable.
3. All Lebesgue integrable functions are integrable in the sense of this chapter.
4. For bounded functions the Lebesgue integral and the integral of this chapter are completely equivalent.

5. For nonnegative functions the Lebesgue integral and the integral of this chapter are completely equivalent.
6. The Lebesgue integral is an absolute integral, i.e., if f is integrable then so too is the absolute value $|f|$.
7. The Lebesgue integral is considered the most important integration theory on the real line and yet viewed as too difficult for most undergraduate mathematics students. (?)

Further study of the Lebesgue integral requires learning the measure theory. The traditional approach is to start with the measure theory and arrive at these descriptive descriptions of his integral only after many weeks. There is an abundance of good texts for this. Try to remember when you are going through such a study that eventually, after much detail, you will indeed arrive back at this point of seeing the integral as an antiderivative.

4.12 The Riemann integral

The last word in Part I of our text goes to the unfortunate Riemann integral, long taught to freshman calculus students in spite of the clamor against it. The formal definition is familiar, of course, since we have already studied the notion of uniform approximation by Riemann sums in Section 3.3.2.

The Riemann integral does not go “beyond the calculus integral.” The Riemann integral will handle no unbounded functions and we have been successful with the calculus integral in handling many such functions. Even for bounded functions the relation between the calculus integral and the Riemann integral is confused: there are functions integrable in either of these senses, but not in the other.

Definition 4.25 Let f be a bounded function that is defined at every point of $[a, b]$. Then, f is said to be Riemann integrable on $[a, b]$ if there is a number I so that for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$\left| I - \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \right| < \varepsilon$$

whenever $\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\}$ is a partition of $[a, b]$ with each

$$x_i - x_{i-1} < \delta \quad \text{and} \quad \xi_i \in [x_{i-1}, x_i].$$

The number I is set equal to $\int_a^b f(x) dx$ and the latter is called the Riemann integral of f on $[a, b]$.

Here are some remarks that you should be able to prove or research.

1. There are Riemann integrable functions that are not integrable [calculus sense].
2. There are bounded, integrable functions [calculus sense] that are not Riemann integrable.
3. All Riemann integrable functions are integrable in the sense of this chapter.
4. All Riemann integrable functions are integrable in the sense of Lebesgue.
5. A bounded function is Riemann integrable if and only if it is continuous at every point, excepting possibly at points in a set of measure zero.
6. The Riemann integral is considered to be a completely inadequate theory of integration and yet is the theory that is taught to most undergraduate mathematics students. (?)

We do not believe that you need to know more about the Riemann integral than these bare facts. Certainly any study that starts with Definition 4.25 and attempts to build and prove a theory of integration is a waste of time; few of the techniques generalize to other settings.

Part II

Theory of the Integral on the Real Line

Chapter 5

Covering Theorems

We embark now on a complete theory for the integral on the real line. In Chapter 3 we studied a very narrow integration theory, which we will now call the *naive calculus integral*, describable as simply interpreting the statement

$$\int_a^b F'(x) dx = F(b) - F(a)$$

to require that $F : [a, b] \rightarrow \mathbb{R}$ is uniformly continuous and differentiable mostly everywhere.¹

In Chapter 4 we extended this definition by relaxing “mostly everywhere” to almost everywhere.² To investigate this new extended calculus integral requires some further techniques before we can advance much further than we did in Chapter 4.

There are a few highlights of the material from Chapter 4 that we wish the reader to recall (or study). We will go deeper now by exploring measure zero sets in greater detail. In particular we prove the Mini-Vitali covering theorem that characterizes measure zero sets in terms of full and fine covers.

Here is our goal for both the review and the new material that will be introduced in this chapter:

- covering relations.

¹Mostly everywhere = everywhere with at most finitely many exceptions.

²Almost everywhere = everywhere excepting a set of measure zero.

- Riemann sums.
- measure zero set.
- full null set.
- fine null set.
- Mini-Vitali theorem asserting the equivalence of measure zero, full null, and fine null.
- zero variation and its relation to zero derivative.
- absolute continuity.
- Lebesgue differentiation theorem asserting the almost everywhere differentiability of functions of bounded variation.

5.1 Covering Relations

The language of integration theory and many of our most important techniques, as presented in the next few chapters, depends on an understanding of and facility with partitions and Riemann sums. A partition is a special case of a covering relation. This section defines and reviews all of the terminology and examines all of the techniques needed to carry on to a complete investigation of the integral.

5.1.1 Partitions and subpartitions

Construct a subdivision of a compact interval $[a, b]$ by choosing points

$$a = a_0 < a_1 < a_2 < \cdots < a_{k-1} < a_k = b$$

and then select points $\xi_1, \xi_2, \dots, \xi_k$ so that each point ξ_i belongs to the corresponding interval $[a_{i-1}, a_i]$. Then the collection

$$\pi = \{([a_{i-1}, a_i], \xi_i) : i = 1, 2, \dots, k\}$$

is called a *partition* of $[a, b]$. Note that the intervals do not overlap and that their union is the whole of the interval $[a, b]$. The associated points must be selected from their corresponding interval. Any subset of a partition is called a *subpartition*.

We consider this a special kind of *covering relation*.

5.1.2 Covering relations

Families of pairs $([u, v], w)$, where $[u, v]$ is a compact interval and w a point in that interval, are called *covering relations*. Every partition and every subpartition is a covering relation. It is a *relation* because it provides an association of points with intervals.

All covering relations are just subsets of one big covering relation:

$$\hat{\beta} = \{([u, v], w) : u, v, w \in \mathbb{R}, u < v \text{ and } u \leq w \leq v\}.$$

We shall most frequently use the Greek symbol β to denote a covering relation. We have already used the Greek symbol π to denote those covering relations which are partitions or subpartitions.

5.1.3 Prunings

Given a number of covering relations arising in a problem we often have to combine them or “prune out” certain subsets of them. We use the following techniques quite frequently:

Definition 5.1 *If β is a covering relation and E a set of real numbers then we write:*

- $\beta[E] = \{([u, v], w) \in \beta : w \in E\}.$
- $\beta(E) = \{([u, v], w) \in \beta : [u, v] \subset E\}.$

to indicate these subsets of the covering relation β from which we have removed inconvenient members.

5.1.4 Full covers

A full cover is one that, in very loose language, contains all sufficiently small intervals at a point.

Definition 5.2 Let E be a set of real numbers. A covering relation β is said to be a full cover of E if for each $w \in E$ there is a positive number $\delta(w)$ so that β contains every pair $([u, v], w)$ for which $v - u < \delta(w)$.

By a *full cover* without reference to any set we mean a full cover of all of $\mathbb{R}$.

Full covers arise naturally as ways to describe continuity, differentiation, integration, and numerous other processes of analysis. The student should attempt many (perhaps all) of the exercises in order to gain a facility in covering arguments.

5.1.5 Fine covers

A fine cover³ is one that, in very loose language, contains arbitrarily small intervals at a point.

Definition 5.3 Let E be a set of real numbers. A covering relation β is said to be a fine cover of E if for each $w \in E$ and any positive number ε the covering relation β contains at least one pair $([u, v], w)$ for which $v - u < \varepsilon$.

By a *fine cover* without reference to any set we mean a fine cover of all of $\mathbb{R}$.

Fine covers arise in the same way that full covers arise. In a sense the fine cover comes from a negation of a full cover. For example (as you will see in the Exercises) full covers could be used to describe continuity conditions and fine covers would then twist this to describe the situation where continuity fails.

5.1.6 Uniformly full covers

A uniformly full cover is one that, in very loose language, contains all sufficiently small intervals at a point, where the smallness required is considered the same for all points

³Known also as a *Vitali cover*.

Definition 5.4 Let E be a set of real numbers. A covering relation β is said to be a uniformly full cover of E if there is a positive number δ so that β contains every pair $([u, v], w)$ for which $v - u < \delta$.

Only occasionally shall we use uniformly full covers. To verify that a covering relation is full just requires us to test what happens at each point. To verify that a covering relation is *uniformly* full requires more: we have to find a positive number δ that works at every point. The exclusive use of uniformly full covers would lead to a restrictive theory: the Riemann integral (which is mostly banished from this textbook) is based on uniformly full covers. Our integration theory uses full covers and, as a consequence, is much more general and is easier.⁴

Exercises

Exercise 534 Suppose that G is an open set. Show that

$$\beta = \{([u, v], w) : u \leq w \leq v, [u, v] \subset G\}$$

is a full cover of G . □

Exercise 535 Suppose that β is a full cover of a set E and that G is an open set containing E . Show that $\beta(G)$ is also a full cover of E . [This is described as “pruning the full cover” by the open set G .] Answer □

Exercise 536 Suppose that β is a fine cover of a set E and that G is an open set containing E . Show that $\beta(G)$ is also a fine cover of E . [This is described as “pruning the fine cover” by the open set G .] Answer □

Exercise 537 Suppose that β is a uniformly full cover of a set E and that G is an open set containing E . Show that $\beta(G)$ is not necessarily a uniformly full cover of E . Would it be a full cover? □

Exercise 538 Suppose that β_1 and β_2 are both full covers of a set E . Show that $\beta_1 \cap \beta_2$ is also a full cover of E . □

Exercise 539 Suppose that β_1 and β_2 are both fine covers of a set E . Show that $\beta_1 \cap \beta_2$ need not be a fine cover of E . □

⁴It is easier since the requirement in Riemann integration to always check that the covers used are not merely full, but uniformly full, imposes unnecessary burdens on many proofs.

Exercise 540 Suppose that β_1 is a full cover of a set E and β_2 is a fine cover. Show that $\beta_1 \cap \beta_2$ is also a fine cover of E .
Need it be a full cover? □

Exercise 541 Suppose that β_1 and β_2 are full covers of sets E_1 and E_2 respectively. Show that $\beta_1 \cup \beta_2$ is a full cover of $E_1 \cup E_2$. □

Exercise 542 Suppose that β_1 and β_2 are fine covers of sets E_1 and E_2 respectively. Show that $\beta_1 \cup \beta_2$ is a fine cover of $E_1 \cup E_2$. □

Exercise 543 Let $E_1, E_2, E_3, \dots$ be a sequence of sets. Suppose that $\beta_1, \beta_2, \beta_3, \dots$ are full covers of sets $E_1, E_2, E_3, \dots$ respectively. Show that

$$\beta = \beta_1 \cup \beta_2 \cup \beta_3 \cup \dots$$

is a full cover of $E = \bigcup_{n=1}^{\infty} E_n$. □

Exercise 544 Let $E_1, E_2, E_3, \dots$ be a sequence of sets. Suppose that $\beta_1, \beta_2, \beta_3, \dots$ are fine covers of sets $E_1, E_2, E_3, \dots$ respectively. Show that

$$\beta = \beta_1 \cup \beta_2 \cup \beta_3 \cup \dots$$

is a fine cover of $E = \bigcup_{n=1}^{\infty} E_n$. □

Exercise 545 Let $F : \mathbb{R} \rightarrow \mathbb{R}$. Define

$$\beta = \{([u, v], w) : |F(u) - F(v)| < \varepsilon\}.$$

Show that β is full at a point x_0 for all $\varepsilon > 0$ if and only if F is continuous at that point. □

Exercise 546 Let $F : \mathbb{R} \rightarrow \mathbb{R}$, $c \in \mathbb{R}$ and define

$$\beta = \{([u, v], w) : |F(u) - F(v) - c(v - u)| < \varepsilon(v - u)\}.$$

Show that β is full at a point x_0 for all $\varepsilon > 0$ if and only if $F'(x_0) = c$. □

Exercise 547 Let $F : \mathbb{R} \rightarrow \mathbb{R}$ and define

$$\beta = \{([u, v], w) : |F(u) - F(v)| \geq \varepsilon\}.$$

Show that β is fine at a point x_0 for some value of $\varepsilon > 0$ if and only if F is not continuous at that point. □

Exercise 548 Let $F : \mathbb{R} \rightarrow \mathbb{R}$, $c \in \mathbb{R}$ and define

$$\beta = \{([u, v], w) : |F(u) - F(v) - c(v - u)| \geq \varepsilon(v - u)\}.$$

Show that β is fine at a point x_0 for some value of $\varepsilon > 0$ if and only if $F'(x_0) = c$ is false. □

Exercise 549 Show that β is fine at a point w if and only if for all β_1 that are full at w there is at least one pair $([u, v], w)$ belonging to both β and β_1 . Answer □

Exercise 550 Show that β is full at a point w if and only if for all β_1 that are fine at w there is at least one pair $([u, v], w)$ belonging to both β and β_1 . Answer □

Exercise 551 (Heine-Borel) Let $\mathcal{G}$ be a family of open sets so that every point in a compact set K is contained in at least one member of the family. Show that the covering relation

$$\beta = \{(I, x) : x \in I \text{ and } I \subset G \text{ for some } G \in \mathcal{G}\}.$$

is a full cover of K (cf. the Heine-Borel Theorem). □

Exercise 552 (Bolzano-Weierstrass) Let E be an infinite set that contains no points of accumulation. Show that

$$\beta = \{(I, x) : x \in I \text{ and } I \cap E \text{ is finite}\}.$$

must be a full cover (cf. the Bolzano-Weierstrass Theorem). □

Exercise 553 Let $\{x_n\}$ be a sequence of real numbers and let

$$\beta = \{(I, x) : x \in I \text{ and } I \text{ contains only finitely many of the } x_n\}.$$

If β is a fine cover of a set E what (if anything) can you conclude? Answer □

Exercise 554 Let $\{x_n\}$ be a sequence of real numbers and let

$$\beta = \{(I, x) : x \in I \text{ and } I \text{ contains only finitely many of the } x_n\}.$$

If β is not a fine cover of a set E what (if anything) can you conclude? Answer □

Exercise 555 Let $\{x_n\}$ be a sequence of real numbers and let

$$\beta = \{(I, x) : x \in I \text{ and } I \text{ contains only finitely many of the } x_n\}.$$

If β is a full cover of a set E what (if anything) can you conclude? Answer □

Exercise 556 Let $\{x_n\}$ be a sequence of real numbers and let

$$\beta = \{(I, x) : x \in I \text{ and } I \text{ contains only finitely many of the } x_n\}.$$

If β is not a full cover of a set E what (if anything) can you conclude?

Answer ☐

Exercise 557 Let $\{x_n\}$ be a sequence of real numbers and let

$$\beta = \{(I, x) : x \in I \text{ and } I \text{ contains infinitely many of the } x_n\}.$$

If β is a fine cover of a set E what (if anything) can you conclude?

Answer ☐

Exercise 558 Let $\{x_n\}$ be a sequence of real numbers and let

$$\beta = \{(I, x) : x \in I \text{ and } I \text{ contains infinitely many of the } x_n\}.$$

If β is not a fine cover of a set E what (if anything) can you conclude?

Answer ☐

Exercise 559 Let $\{x_n\}$ be a sequence of real numbers and let

$$\beta = \{(I, x) : x \in I \text{ and } I \text{ contains infinitely many of the } x_n\}.$$

If β is a full cover of a set E what (if anything) can you conclude?

Answer ☐

Exercise 560 Let $\{x_n\}$ be a sequence of real numbers and let

$$\beta = \{(I, x) : x \in I \text{ and } I \text{ contains infinitely many of the } x_n\}.$$

If β is not a full cover of a set E what (if anything) can you conclude?

Answer ☐

5.1.7 Cousin covering lemma

Throughout Chapters 1–4 we have made extensive use of the Cousin covering lemma, but we repeat it here for convenience and to stress the role that it plays in covering arguments in analysis and in integration theory. This also allows us a chance to rewrite the proof in the language of this chapter.

Lemma 5.5 (Cousin covering lemma) *Let β be a full cover. Then β contains a partition of every compact interval.*

Proof. Note, first, that if β fails to contain a partition of some interval $[a, b]$ then it must fail to contain a partition of much smaller subintervals. For example if $a < c < b$, if π_1 is a partition of $[a, c]$ and π_2 is a partition of $[c, b]$, then $\pi_1 \cup \pi_2$ is certainly a partition of $[a, b]$.

We use this feature repeatedly. Suppose that β fails to contain a partition of $[a, b]$. Choose a subinterval $[a_1, b_1]$ with length less than $1/2$ the length of $[a, b]$ so that β fails to contain a partition of $[a_1, b_1]$. Continue inductively, selecting a nested sequence of compact intervals $[a_n, b_n]$ with lengths shrinking to zero so that β fails to contain a partition of each $[a_n, b_n]$.

By the nested interval property there is point z belonging to each of the intervals. As β is a full cover, there must exist a $\delta > 0$ so that β contains (I, z) for any compact subinterval I of $[a, b]$ with length smaller than δ . In particular β contains all $([a_n, b_n], z)$ for n large enough to assure us that $b_n - a_n < \delta$. The set $\pi = \{([a_n, b_n], z)\}$ containing a single element is itself a partition of $[a_n, b_n]$ that is contained in β . That contradicts our assumptions. Consequently β must contain a partition of $[a, b]$. Since $[a, b]$ was arbitrary, β must contain a partition of any compact interval. ■

5.1.8 Decomposition of full covers

There is a decomposition of full covers that is often of use in constructing a proof. Here is a good place to put it for easy reference, although it is entirely unmotivated for the moment. This shows that, while a full cover is a much more general object than a uniformly full cover, it can be broken into pieces that are themselves uniform covers.

Lemma 5.6 (Decomposition Lemma) *Let β be a full cover of a set E . Then there is an increasing sequence of sets $\{E_n\}$ with $E = \bigcup_{n=1}^{\infty} E_n$ and a sequence of nonoverlapping compact intervals $\{I_{kn}\}$ covering E_n so that if x is any point in E_n and I is any subinterval of I_{kn} that contains x then (I, x) belongs to β .*

Proof. Let β be a full cover of a set E . By the nature of the cover there must exist, for each $x \in E$ a positive number $\delta(x)$ on E with the property that (I, x) belongs to β whenever if $x \in E$, $x \in I$ and the length of the interval I is smaller than $\delta(x)$. Define

$$E_n = \{x \in E : \delta(x) > 1/n\}.$$

This is an expanding sequence of subsets of E whose union is E itself. If I is any compact interval that contains a point x in E_n and has length less than $1/n$, then (I, x) must belong to β .

A way of exploiting this property is to introduce the intervals

$$I_{mn} = \left[\frac{m}{n}, \frac{m+1}{n} \right]$$

for integers $m = 0, \pm 1, \pm 2, \dots$. Then $\beta([E_n \cap I_{mn}])$ has this property: if x is any point in $E_n \cap I_{mn}$ and I is any subinterval of I_{mn} that contains x then (I, x) is a member of $\beta([E_n \cap I_{mn}])$.

Thus the condition of being a full cover, which is a local condition defined in a special way at each point, has been made uniform throughout each piece of the decomposition. If we relabel these sets in a convenient way then we now have our decomposition property. ■

5.1.9 Riemann sums

The integral can be characterized as a limit of Riemann sums. The original Riemann integral has such a definition and the Lebesgue integral, although defined in a completely different manner, also has such a characterization although not as simple as that for the Riemann integral.

In fact we will wish to define upper and lower integrals, so the upper integral is a limsup of Riemann sums and the lower integral is a liminf of Riemann sums. The notation for Riemann sums can assume any of the following forms (5.1), (5.2), (5.3), or (5.4), depending on which is convenient:

Take an interval $[a, b]$ and subdivide as follows:

$$a = x_0 < x_1 < x_2 < x_3 < \dots < x_{n-1} < x_n = b.$$

Then form a partition of $[a, b]$ by selecting points ξ_i from each of the corresponding intervals:

$$\pi = ([x_0, x_1], \xi_1), ([x_1, x_2], \xi_2), \dots, ([x_{n-1}, x_n], \xi_n).$$

Sums of the following form are called *Riemann sums* with respect to this partition:

$$\sum_{k=1}^n f(\xi_k)(x_k - x_{k-1}). \quad (5.1)$$

These can also be more conveniently written as

$$\sum_{([u, v], w) \in \pi} f(w)(v - u) \quad (5.2)$$

5.1. COVERING RELATIONS

197

or

$$\sum_{([u,v],w) \in \pi} f(w) \lambda([u,v]) \quad (5.3)$$

or even as

$$\sum_{(I,w) \in \pi} f(w) \lambda(I). \quad (5.4)$$

Here we are using λ as a *length function*:

$$\lambda([u,v]) = v - u$$

is simply the length of the interval $[u,v]$. We can in this way also conveniently assign a length to the intersection of two compact intervals. For example,

$$\lambda([u_1, v_1] \cap [u_2, v_2])$$

would be the length of the interval $[u_1, v_1] \cap [u_2, v_2]$ (if it is an interval) and would have length zero if the two intervals do not overlap.

General Riemann sums In general, let $h([u,v], w)$ denote any real-valued function which assigns to an interval-point pair $([u,v], w)$ a real value. Let π be any partition or subpartition. Then we will (loosely) call any sum

$$\sum_{([u,v],w) \in \pi} h([u,v], w) \quad (5.5)$$

or

$$\sum_{(I,w) \in \pi} h(I, w) \quad (5.6)$$

a Riemann sum. Such sums will play a role in many diverse investigations.

Exercises

Exercise 561 Let $F : [a, b] \rightarrow \mathbb{R}$ and let π be a partition of $[a, b]$. Verify the computations

$$\sum_{([u,v],w) \in \pi} (v - u) = b - a$$

and

$$\sum_{([u,v],w) \in \pi} (F(v) - F(u)) = F(b) - F(a).$$

□

Exercise 562 Let $F : [a, b] \rightarrow \mathbb{R}$ and let π be a partition of $[a, b]$. Show that

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)| \geq |F(b) - F(a)|.$$

□

Exercise 563 Let $F : [a, b] \rightarrow \mathbb{R}$ be a Lipschitz function with Lipschitz constant M and let π be a partition of the interval $[a, b]$. Show that

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)| \leq M(b - a).$$

□

Exercise 564 Let $F, f : [a, b] \rightarrow \mathbb{R}$ and let π be a partition of $[a, b]$ and suppose that

$$F(v) - F(u) \geq f(w)(v - u)$$

for all $([u, v], w) \in \pi$. Show that

$$\sum_{([u,v],w) \in \pi} f(w)(v - u) \leq F(b) - F(a).$$

□

Exercise 565 Let $F : [a, b] \rightarrow \mathbb{R}$ be a function with the property that

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)| = 0.$$

for every partition π of the interval $[a, b]$. What can you conclude?

□

Exercise 566 Let $F : [0, 1] \rightarrow \mathbb{R}$ be a function with the property that it is monotonic on each of the intervals $[0, \frac{1}{3}]$, $[\frac{1}{3}, \frac{2}{3}]$,

5.1. COVERING RELATIONS

and $[\frac{2}{3}, 1]$. What is the largest possible value of

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)|$$

for arbitrary partitions π of the interval $[a, b]$. □

Exercise 567 Describe the difference between the two sums

$$\sum_{([u,v],w) \in \pi} f(w)(v - u)$$

and

$$\sum_{(I,w) \in \pi([c,d])} f(w)(v - u)$$

where $[c, d]$ is an interval. Answer □

Exercise 568 Describe the difference between the two sums

$$\sum_{([u,v],w) \in \pi} f(w)(v - u)$$

and

$$\sum_{([u,v],w) \in \pi[E]} f(w)(v - u).$$

where E is a set. Answer □

Exercise 569 How could you interpret the expression

$$\sum_{([u,v],w) \in \pi_1 \cup \pi_2} f(w)(v - u)?$$

□

Exercise 570 How could you interpret the expression

$$\sum_{([u_1,v_1],w_1) \in \pi_1} \sum_{([u_2,v_2],w_2) \in \pi_2} f(w_1) \lambda([u_1, v_1] \cap [u_2, v_2])?$$

if π_1 and π_2 are both partitions of the same interval $[a, b]$? □

Exercise 571 Show that

$$\begin{aligned} & \sum_{([u_1, v_1], w_1) \in \pi_1} f(w_1) \lambda([u_1, v_1]) - \sum_{([u_2, v_2], w_2) \in \pi_2} f(w_2) \lambda([u_2, v_2]) = \\ & \sum_{([u_1, v_1], w_1) \in \pi_1} \sum_{([u_2, v_2], w_2) \in \pi_2} [f(w_1) - f(w_2)] \lambda([u_1, v_1] \cap [u_2, v_2]) \end{aligned}$$

if π_1 and π_2 are both partitions of the same interval $[a, b]$? □

Exercise 572 Let $f : [a, b] \rightarrow \mathbb{R}$ be a continuous function. What could you require of two partitions π_1 and π_2 of the interval $[a, b]$ in order to conclude that

$$\left| \sum_{([u_1, v_1], w_1) \in \pi_1} f(w_1)(v_1 - u_1) - \sum_{([u_2, v_2], w_2) \in \pi_2} f(w_2)(v_2 - u_2) \right| < \varepsilon?$$

□

5.2 Sets of measure zero

We review the notion of a set of measure zero already studied in Chapter 4. We will present three distinct versions of measure zero. The first is due to Lebesgue and arises from his theory of measure. The second and third use full and fine coverings and estimates using Riemann sums. In Chapter 4 we used the full covering version for our first definition of measure zero. Now we begin with Lebesgue's definition.

5.2.1 Lebesgue measure of open sets

The property that a set E will be a set of measure zero is actually a statement about the family of open sets containing E . A set E is measure zero if there are arbitrarily “small” open sets containing E .

For a precise version of this we require a definition for the Lebesgue measure $\lambda(G)$ of an open set G . Later on, in Chapter 7, we will study Lebesgue's measure in general. The attention here is directed only on that measure for open sets.

Definition 5.7 Let G be an open set. Then the Lebesgue measure $\lambda(G)$ of an open set G is defined to be the total sum of the lengths of all the component intervals of G .

According to this definition $\lambda(\emptyset) = 0$ (since there are no component intervals). If G consists of infinitely many bounded component intervals $(\{a_i, b_i\})$ then the measure is the sum of an infinite series:

$$\lambda(G) = \sum_{i=1}^{\infty} (b_i - a_i).$$

[An unbounded component interval would have length ∞ and so an open set with an unbounded component has infinite measure.]

The only tool we need for working with this concept for the moment is given by the subadditivity property.

Lemma 5.8 (Subadditivity) Let $G_1, G_2, G_3, \dots$ be a sequence of open sets. Then the union

$$G = \bigcup_{i=1}^{\infty} G_i$$

is also an open set and

$$\lambda(G) \leq \sum_{i=1}^{\infty} \lambda(G_i).$$

Proof. Certainly G is open since any union of open sets is open. Let

$$T = \sum_{i=1}^{\infty} \lambda(G_i).$$

Note that T is simply the sum of the lengths of all the component intervals of all the G_i .

Let $(\{a_j, b_j\})$ denote the component intervals of G . Take (a_1, b_1) and choose any $[c_1, d_1] \subset (a_1, b_1)$. A compactness argument shows that $[c_1, d_1]$ is contained in finitely many of the component intervals making up the sum T . We conclude that $d_1 - c_1 \leq T$. That would be true for any choice of $[c_1, d_1] \subset (a_1, b_1)$, so that $b_1 - a_1 \leq T$. A similar argument using m components $(a_1, b_1), (a_2, b_2), \dots, (a_m, b_m)$ will establish that

$$\sum_{j=1}^m (b_j - a_j) \leq T$$

from which

$$\lambda(G) = \sum_{j=1}^{\infty} (b_j - a_j) \leq T$$

evidently follows. ■

5.2.2 Sets of Lebesgue measure zero

Our first definition of measure zero set expresses this as a property of open sets that contain the set.

Definition 5.9 *Let E be a set of real numbers. Then E is said to have measure zero if for every $\varepsilon > 0$ there is an open set G containing E for which $\lambda(G) < \varepsilon$.*

Recall that we have given a completely different definition of measure zero in Chapter 4. Thus we are obliged very quickly to show that these two definitions are equivalent. In the meantime the following exercises should be attempted but now with the new definition. In Section 5.5 we will show that the two definitions (along with a third definition for measure zero) are equivalent.

Exercises

Exercise 573 *The empty set has measure zero.*

Answer ☐

Exercise 574 *Every finite set has measure zero.*

Answer ☐

Exercise 575 *Every infinite, countable set has measure zero.*

Answer ☐

Exercise 576 *The Cantor set has measure zero.*

Answer ☐

5.2.3 Sequences of measure zero sets

One of the most fundamental of the properties of sets having measure zero is how sequences of such sets combine. We recall that the union of any sequence of countable sets is also countable. We now prove that the union of any sequence of measure zero sets is also a measure zero set.

Theorem 5.10 *Let $E_1, E_2, E_3, \dots$ be a sequence of sets of measure zero. Then the set E formed by taking the union of all the sets in the sequence is also of measure zero.*

Proof. Let $\varepsilon > 0$. Choose open sets $G_n \supset E_n$ so that

$$\lambda(G_n) < 2^{-n}\varepsilon.$$

Then set $G = \bigcup_{n=1}^{\infty} G_n$. Observe, by the subadditivity property (i.e., from Lemma 5.8), that G is an open set containing E for which $\lambda(G) < \varepsilon$. ■

Exercises

Exercise 577 *Show that E is a set of measure zero if and only if there is a finite or infinite sequence*

$$(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4), \dots$$

of open intervals covering the set E so that

$$\sum_{k=1}^{\infty} (b_k - a_k) \leq \varepsilon.$$

□

Exercise 578 (compact sets of measure zero) *Let E be a compact set of measure zero. Show that for every $\varepsilon > 0$ there is a finite collection of open intervals*

$$\{(a_k, b_k) : k = 1, 2, 3, \dots, N\}$$

that covers the set E and so that

$$\sum_{k=1}^N (b_k - a_k) < \varepsilon.$$

Answer □

Exercise 579 *Show that E is a set of measure zero if and only if there is a finite or infinite sequence*

$$[a_1, b_1], [a_2, b_2], [a_3, b_3], [a_4, b_4], \dots$$

of compact intervals covering the set E so that

$$\sum_{k=1}^{\infty} (b_k - a_k) \leq \varepsilon.$$

□

Exercise 580 Show that every subset of a set of measure zero also has measure zero.

□

Exercise 581 Suppose that $E \subset [a, b]$ is a set of measure zero. Show that $\int_a^b \chi_E(x) dx = 0$.

□

Exercise 582 If E has measure zero, show that the translated set

$$E + \alpha = \{x + \alpha : x \in E\}$$

also has measure zero.

□

Exercise 583 If E has measure zero, show that the expanded set

$$cE = \{cx : x \in E\}$$

also has measure zero for any $c > 0$.

□

Exercise 584 If E has measure zero, show that the reflected set

$$-E = \{-x : x \in E\}$$

also has measure zero.

□

Exercise 585 Without referring to Theorem 5.10, show that the union of any two sets of measure zero also has measure zero.

□

Exercise 586 If $E_1 \subset E_2$ and E_1 has measure zero but E_2 has not, what can you say about the set $E_2 \setminus E_1$?

□

Exercise 587 Show that any interval (a, b) or $[a, b]$ is not of measure zero.

□

Exercise 588 Give an example of a set that is not of measure zero and does not contain any interval $[a, b]$.

□

5.2. SETS OF MEASURE ZERO

205

Exercise 589 A careless student claims that if a set E has measure zero, then it is true that the closure $\overline{E}$ must also have measure zero. After all, if E is contained in $\bigcup_{i=1}^{\infty} (a_i, b_i)$ with small total length then $\overline{E}$ is contained in $\bigcup_{i=1}^{\infty} [a_i, b_i]$, also with small total length. Is this correct? □

Exercise 590 If a set E has measure zero what can you say about interior points of that set? □

Exercise 591 If a set E has measure zero what can you say about boundary points of that set? □

Exercise 592 Suppose that a set E has the property that $E \cap [a, b]$ has measure zero for every compact interval $[a, b]$. Must E also have measure zero? □

Exercise 593 Show that the set of real numbers in the interval $[0, 1]$ that do not have a 7 in their infinite decimal expansion is of measure zero. □

Exercise 594 Describe completely the class of sets E with the following property: For every $\varepsilon > 0$ there is a finite collection of open intervals

$$(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4), \dots, (a_N, b_N)$$

that covers the set E and so that

$$\sum_{k=1}^N (b_k - a_k) < \varepsilon.$$

(These sets are said to have zero content.) □

Exercise 595 Show that a set E has measure zero if and only if there is a sequence of intervals

$$(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4), \dots$$

so that every point in E belongs to infinitely many of the intervals and $\sum_{k=1}^{\infty} (b_k - a_k)$ converges. □

Exercise 596 Suppose that $\{(a_i, b_i)\}$ is a sequence of open intervals for which

$$\sum_{i=1}^{\infty} (b_i - a_i) < \infty.$$

Show that the set

$$E = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} (a_i, b_i)$$

has measure zero. What relation does this exercise have with the preceding exercise? □

Exercise 597 By altering the construction of the Cantor set, construct a nowhere dense closed subset of $[0, 1]$ so that the sum of the lengths of the intervals removed is not equal to 1. Will this set have measure zero? □

5.2.4 Almost everywhere language

Some commonly used language is used in discussions of measure zero sets. Let $P(x)$ be a property that may or not be possessed by a point $x \in \mathbb{R}$. We say that

$P(x)$ is true almost everywhere.

or

$P(x)$ is true for almost every x .

if the set

$$\{x \in \mathbb{R} : P(x) \text{ is not true}\}$$

is a measure zero set.

We have suggested this language before in Section 2.1.1 and we shall now officially take it on. Thus we write:

(mostly everywhere) A statement holds *mostly everywhere* if it holds everywhere with the exception of a finite set of points $c_1, c_2, c_3, \dots, c_n$.

(nearly everywhere) A statement holds *nearly everywhere* if it holds everywhere with the exception of a countable set.

(almost everywhere) A statement holds *almost everywhere* if it holds everywhere with the exception of a set of measure zero.

Nearly everywhere might be abbreviated “n.e.” but only in a context where the reader is reminded of such usage. Almost everywhere is very frequently abbreviated “a.e.” and most advanced readers are familiar with this usage.

Exercises

Exercise 598 What would it mean to say that a function is almost everywhere discontinuous? □

Exercise 599 What would it mean to say that a function is almost everywhere differentiable? Give an example of function that is almost everywhere differentiable, but not everywhere differentiable. □

Exercise 600 What would it mean to say that almost every point in $\mathbb{R}$ is irrational? Is this a true statement? □

Exercise 601 What would it mean to say that almost everywhere point in a set A belongs to a set B ? Give an example for which this is true and an example for which this is false. □

Exercise 602 What would it mean to say that a function is almost everywhere equal to zero? □

Exercise 603 What would it mean to say that a function is almost everywhere different from zero? □

Exercise 604 Suppose that the function $f : [a, b] \rightarrow \mathbb{R}$ is integrable and is almost everywhere in $[a, b]$ nonnegative. Show that $\int_a^b f(x) dx \geq 0$. □

Exercise 605 Suppose that the functions $f, g : [a, b] \rightarrow \mathbb{R}$ are integrable and that $f(x) \leq g(x)$ for almost every x in $[a, b]$. Show that $\int_a^b f(x) dx \leq \int_a^b g(x) dx$. □

Exercise 606 Suppose that the functions $F, G : [a, b] \rightarrow \mathbb{R}$ are continuous almost everywhere in $[a, b]$. Is the sum function $F(x) + G(x)$ also continuous almost everywhere in $[a, b]$. □

Exercise 607 Suppose that the functions $F, G : [a, b] \rightarrow \mathbb{R}$ are differentiable almost everywhere in $[a, b]$. Is the sum function $F(x) + G(x)$ also differentiable almost everywhere in $[a, b]$. □

5.3 Full null sets

Sets of measure zero are defined using open sets that contain them. There is a variant on this using full covers instead. We have already taken advantage of this variant in Chapter 4 because that variant has the closest connection with integration

theory as we have presented it. For the moment we refer to this version of measure zero as “full null.” Once we have proved the equivalence we can revert to normal usage and just label such sets as “measure zero.”

This new definition has the advantage that, since it is defined using full covers, the definition is more closely related to the differentiation and integration properties of functions. It has the disadvantage that, unlike the measure zero sets, it is not constructive; full covers themselves are not necessarily constructive.

Definition 5.11 *A set E of real numbers is said to be full null if for every $\varepsilon > 0$ there is a full cover β of the set E with the property that*

$$\sum_{([u,v],w) \in \pi} (v - u) < \varepsilon \quad (5.7)$$

for every subpartition π chosen from β .

We will show that the two definitions, full null and measure zero, are equivalent later. For the moment one direction is easy.

Theorem 5.12 *Every set of measure zero is also full null.*

Proof. Assume that a set E measure zero and let $\varepsilon > 0$. Choose an open set G containing E for which $\lambda(G) < \varepsilon$. Let $\{(a_i, b_i)\}$ be the component intervals of G . Define β to be the collection of all pairs $([u, v], w)$ with the property that $w \in [u, v] \subset G$. It is easy to check that β is a full cover of E .

Consider any subpartition π chosen from β . For each $([u, v], w) \in \pi$, $[u, v]$ is a subinterval of some component (a_i, b_i) of G . Holding i fixed, the sum of the lengths of those intervals $[u, v] \subset (a_i, b_i)$ would certainly be smaller than $(b_i - a_i)$. It follows that

$$\sum_{([u,v],w) \in \pi} (v - u) \leq \sum_{i=1}^{\infty} (b_i - a_i) = \lambda(G) < \varepsilon.$$

This verifies that E is full null. ■

Exercises

Exercise 608 *Show that every subset of a full null set is also a full null set.* □

Exercise 609 Show that $E \subset [a, b]$ is a full null set if and only if $\int_a^b \chi_E(x) dx = 0$. Compare this with Exercise 581. $\square$

Exercise 610 Show that $E \subset \mathbb{R}$ is a full null set if and only if $\int_a^b \chi_E(x) dx = 0$ for every compact interval $[a, b]$. $\square$

Exercise 611 Show that the union of any two full null sets is also a full null set. $\square$

Exercise 612 Show that the union of any sequence of full null sets is also a full null set. $\square$

Exercise 613 Define a set E to be uniformly full null if for every $\varepsilon > 0$ there is a uniformly full cover β of the set E with the property that

$$\sum_{([u,v],w) \in \pi} (v - u) < \varepsilon \quad (5.8)$$

for every subpartition π chosen from β . Show that uniformly full null sets are the same as sets of zero content. (cf. Exercise 594). $\square$

5.4 Fine null sets

Sets of measure zero are defined with attention to the open sets that contain them. Full null sets are defined using full covers. There is a third variant on this using fine covers instead. This offers yet a more delicate way of working with measure zero sets, since fine covers can express very subtle properties of derivatives and integrals. We will show in Section 5.5 that all three notions are equivalent.

Definition 5.13 A set E of real numbers is said to be fine null if for every $\varepsilon > 0$ there is a fine cover β of the set E with the property that

$$\sum_{([u,v],w) \in \pi} (v - u) < \varepsilon \quad (5.9)$$

for every subpartition π chosen from β .

Exercises

Exercise 614 Show that every set of measure zero is also fine null. □

Exercise 615 Show that every full null set is also fine null. □

Exercise 616 Show that every subset of a fine null set is also a fine null set. □

Exercise 617 Show that the union of any two fine null sets is also a full null set. □

Exercise 618 Show that the union of any sequence of fine null sets is also a fine null set. □

5.5 The Mini-Vitali Covering Theorem

The original Vitali covering theorem asserts that the Lebesgue measure of an arbitrary set can be determined either by open coverings of E , or by full covers of E , or by fine covers of E . Our goals in this chapter are narrower. We want to establish these same facts, but only for sets of measure zero. Later, in Chapter 7 we will return and complete the Vitali covering theorem.

Theorem 5.14 (Mini-Vitali covering theorem) For any set $E \subset \mathbb{R}$ the following are equivalent:

1. E is a set of measure zero.
2. E is a full null set.
3. E is a fine null set.

As a result of this theorem we can now simply refer to these sets as measure zero sets and use any of the three characterizations that is convenient. The proof requires some simple geometric arguments and an application of the Heine-Borel theorem; it is given in the sections that now follow.

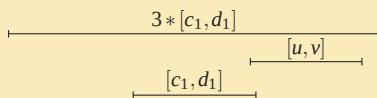


Figure 5.1: Note that $3 * [c_1, d_1]$ will then include any shorter interval $[u, v]$ that intersects $[c_1, d_1]$.

5.5.1 Covering lemmas for families of compact intervals

We begin with some simple covering lemmas for finite and infinite families of compact intervals.

Lemma 5.15 *Let $\mathcal{C}$ be a finite family of compact intervals. Then there is a pairwise disjoint subcollection $[c_i, d_i]$ ($i = 1, 2, \dots, m$) of that family with^a*

$$\bigcup_{[u,v] \in \mathcal{C}} [u, v] \subset \bigcup_{i=1}^m 3 * [c_i, d_i].$$

^aBy $3 * [u, v]$ we mean the interval centered at the same point as $[u, v]$ but with three times the length.

Proof. For $[c_1, d_1]$ simply choose the largest interval. Note that $3 * [c_1, d_1]$ will then include any other interval $[u, v] \in \mathcal{C}$ that intersects $[c_1, d_1]$. See Figure 5.1.

For $[c_2, d_2]$ choose the largest interval from among those that do not intersect $[c_1, d_1]$. Note that together $3 * [c_1, d_1]$ and $3 * [c_2, d_2]$ include any interval of the family that intersects either $[c_1, d_1]$ or $[c_2, d_2]$. Continue inductively, choosing $[c_{k+1}, d_{k+1}]$ as the largest interval in $\mathcal{C}$ that does not intersect one the previously chosen intervals $[c_1, d_1], [c_2, d_2], \dots, [c_k, d_k]$. Stop when you run out of intervals to select. ■

The next covering lemma addresses arbitrary families of compact intervals.

Lemma 5.16 *Let C be any collection of compact intervals. Then the set*

$$G = \bigcup_{[u,v] \in C} (u,v)$$

is an open set that contains all but countably many points of the set

$$E = \bigcup_{[u,v] \in C} [u,v]$$

Proof. Let

$$C = \{x : x \notin G \text{ and } x = c \text{ or } x = d \text{ for at least one } [c,d] \in C\}.$$

We observe that G is open, being a union of a family of open intervals. Clearly G contains all of E except for points that are in the set C . To complete the proof of the lemma, we show that C is countable. Write, for $n = 1, 2, 3, \dots$,

$$C_n = \{x : x \notin G, \ x = c \text{ for at least one } [c,d] \in C \text{ with } d - c > 1/n\}.$$

$$C'_n = \{x : x \notin G, \ x = d \text{ for at least one } [c,d] \in C \text{ with } d - c > 1/n\}.$$

We easily show that each C_n and C'_n is countable. For example if $c \in C_n$ then the interval $(c, c + 1/n)$ can contain no other point of C . This is because there is at least one interval $[c,d]$ from C with $d - c > 1/n$. Thus $(c, c + 1/n) \subset (c,d) \subset G$. Consequently there can be only countably many such points. It follows that the set $C = \bigcup_{n=1}^{\infty} (C_n \cup C'_n)$ is a countable subset of E . ■

5.5.2 Proof of the Mini-Vitali covering theorem

We begin with a simple lemma that is the key to the argument, both for our proof of the mini version as well as the proof of the full Vitali covering theorem.

Lemma 5.17 *Let β be a covering relation and write*

$$G = \bigcup_{([u,v],w) \in \beta} (u,v).$$

Then G is an open set and, if $g = \lambda(G)$, is finite then there must exist a subpartition $\pi \subset \beta$ for which

$$\sum_{([u,v],w) \in \pi} (v-u) \geq g/6. \quad (5.10)$$

In particular

$$G' = G \setminus \bigcup_{([u,v],w) \in \pi} [u,v]$$

is an open subset of G and $\lambda(G') \leq 5g/6$.

Proof. It is clear that the set G of the lemma, expressed as the union of a family of open intervals, must be an open set. Let $\{(a_i, b_i)\}$ be the sequence of component intervals of G . Thus, by definition,

$$g = \lambda(G) = \sum_{i=1}^{\infty} (b_i - a_i).$$

Choose an integer N large enough that

$$\sum_{i=1}^N (b_i - a_i) > 3g/4.$$

Inside each open interval (a_i, b_i) , for $i = 1, 2, \dots, N$, choose a compact interval $[c_i, d_i]$ so that

$$\sum_{i=1}^N (d_i - c_i) > g/2.$$

Write

$$K = \bigcup_{i=1}^N [c_i, d_i]$$

and note that it is a compact set covered by the family

$$\{(u,v) : ([u,v],w) \in \beta\}.$$

By the Heine-Borel theorem there must be a finite subset

$$([u_1, v_1], w_1), ([u_2, v_2], w_2), ([u_3, v_3], w_3), \dots, ([u_m, v_m], w_m)$$

from β for which

$$K \subset \bigcup_{i=1}^m (u_i, v_i).$$

By Lemma 5.15 we can extract a subpartition π from this list so that

$$K \subset \bigcup_{([u,v],w) \in \pi} 3 * [u, v].$$

and so

$$\sum_{([u,v],w) \in \pi} 3(v-u) \geq \sum_{i=1}^N (d_i - c_i) > g/2.$$

Statement (5.10) then follows. [Not that we need it here, but recall that Lemma 5.15 allows us to claim that the intervals in the subpartition π are disjoint, not merely nonoverlapping.]

The final statement of the lemma requires just checking the length of a finite number of the components of G' . We have removed all the intervals $[u, v]$ from G for which $([u, v], w) \in \pi$. Since the total length removed is greater than $g/6$ what remains cannot be larger than $5g/6$. ■

Proof of the Mini-Vitali covering theorem: We already know that every set of measure zero is full null, and that every full null set is fine null. To complete the proof we show that every fine null set is a set of measure zero. Let us suppose that E is not a set of measure zero. We show that it is not fine full then. Define

$$\varepsilon_0 = \inf\{\lambda(G) : G \text{ open and } G \supset E\}.$$

Since E is not measure zero, $\varepsilon_0 > 0$.

Let β be an arbitrary fine cover of E . Define

$$G = \bigcup_{([u,v],w) \in \beta} (u, v).$$

This is an open set and, by Lemma 5.16, G covers all of E except for a countable set. It follows that $\lambda(G) \geq \varepsilon_0$, since

if $\lambda(G) < \varepsilon_0$ we could add to G a small open set G' that contains the missing countable set of points and for which the combined set $G \cup G'$ is an open set containing E but with measure smaller than ε_0 .

By Lemma 5.17 there must exist a subpartition $\pi \subset \beta$ for which

$$\sum_{([u,v],w) \in \pi} (v - u) \geq \varepsilon_0/6.$$

But that means that E is not a fine null set, since this is true for *every* fine cover β .

5.6 Functions having zero variation

A set E is full null (i.e., measure zero) if there is a full cover β of the set E so that

$$\sum_{([u,v],w) \in \pi} (v - u) < \varepsilon$$

whenever π is a subpartition, $\pi \subset \beta$. This generalizes easily by considering instead sums

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)|$$

for some function F . We have used this definition in Chapter 4 but repeat and review it here.

Definition 5.18 Let F be defined on an open set that contains a set E of real numbers. We say that F has zero variation on the set E provided that for every $\varepsilon > 0$ there is a full cover β of the set E so that

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)| < \varepsilon$$

whenever π is a subpartition, $\pi \subset \beta$.

Lemma 5.19 Let $F : (a, b) \rightarrow \mathbb{R}$. Then F has zero variation on the open interval (a, b) if and only if F is constant on (a, b) .

Proof. One direction is obvious; the other direction is an application of the Cousin covering lemma. Suppose that F has

zero variation on (a, b) . Let $\varepsilon > 0$ and choose a full cover β of the set (a, b) so that

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)| < \varepsilon$$

whenever π is a subpartition, $\pi \subset \beta$. If $[c, d] \subset (a, b)$ then there is a partition $\pi \subset \beta$ of the whole interval $[c, d]$. Consequently

$$|F(d) - F(c)| \leq \sum_{([u,v],w) \in \pi} |F(v) - F(u)| < \varepsilon.$$

This holds for every such interval $[c, d]$ and every positive ε . It follows that F must be constant on (a, b) . ■

Lemma 5.20 *Let F be defined on an open set that contains each of the sets $E_1, E_2, E_3, \dots$ and suppose that F has zero variation on each E_i ($i = 1, 2, 3, \dots$). Then F has zero variation on any subset of the union $\bigcup_{i=1}^{\infty} E_i$.*

Proof. Let $\varepsilon > 0$ and, for each integer i , choose a full cover β_i of E_i so that

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)| < 2^{-i}\varepsilon \tag{5.11}$$

whenever π is a subpartition, $\pi \subset \beta_i$. Construct β as the union of the sequence $\beta_i[E_i]$. This is a full cover of any subset E of the union $\bigcup_{i=1}^{\infty} E_i$. Now simply check that, if π is a subpartition, $\pi \subset \beta$ then

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)| \leq \sum_{i=1}^{\infty} \sum_{([u,v],w) \in \pi[E_i]} |F(v) - F(u)| < \sum_{i=1}^{\infty} 2^{-i}\varepsilon = \varepsilon. \tag{5.12}$$

It follows that F has zero variation on E . ■

Exercises

Exercise 619 Show that a constant function has zero variation on any set. □

Exercise 620 Show that if F has zero variation on a set E then it has zero variation on any subset of E . □

Exercise 621 Let E contain a single point x_0 . What does it mean for F to have zero variation on E ?

Answer □

Exercise 622 Let E have countably many points. Show that F has zero variation on the set E if and only if F has zero variation on the singleton sets $\{e\}$ for each $e \in E$. □

Exercise 623 Show that N is a null set if and only if the function $F(x) = x$ has zero variation on N . □

Exercise 624 Suppose that both the functions F and G have zero variation on a set E . Show that so too does every linear combination $rF + sG$. □

Exercise 625 Suppose that both the functions F and G have zero variation on a set E . Does it follow that the product FG must have zero variation on E ? □

Exercise 626 Show that a continuous function has variation zero on every countable set. □

Exercise 627 Show that a function that has variation zero on every countable set must be continuous. □

5.6.1 Zero variation and zero derivatives

There is an intimate connection between the notion of zero variation and the fact of zero derivatives. The following two theorems are central to our theory. Note that zero derivatives imply zero variation and that, conversely, zero variation implies zero derivatives (but only almost everywhere).

Theorem 5.21 Let $F : \mathbb{R} \rightarrow \mathbb{R}$ and suppose that $F'(x) = 0$ at every point of the set E . Then F has zero variation on E .

Proof. Fix an integer n and write $E_n = (-n, n) \cap E$. Let $\varepsilon > 0$ and consider the collection

$$\beta = \{([u, v], w) : w \in E, w \in [u, v] \subset (-n, n), |F(v) - F(u)| < \varepsilon(v - u)\}.$$

By our assumption that $F'(x) = 0$ at every point of E we see easily that β is a full cover of E_n . But if $\pi \subset \beta$ is any subpartition we must have

$$\sum_{([u, v], w) \in \pi} |F(v) - F(u)| < \sum_{([u, v], w) \in \pi} \varepsilon(v - u) < 2\varepsilon n.$$

This proves that F has zero variation on each set E_n . It follows from Lemma 5.20 that F has zero variation on the set E which is, evidently, the union of the sequence of sets $\{E_n\}$. ■

Theorem 5.22 *Let $F : \mathbb{R} \rightarrow \mathbb{R}$ and suppose that F has zero variation on a set E . Then $F'(x) = 0$ at almost every point of the set E .*

Proof. This theorem is deeper than the preceding and will require, for us, an appeal to our version of the Vitali covering theorem. Let N be the set of points x in E at which $F'(x) = 0$ is false. A fine covering argument allows us to analyze this. There must be some positive number $\varepsilon(x)$ for each $x \in N$ so that

$$\beta_1 = \{([u, v], w) : w \in E, |F(v) - F(u)| \geq \varepsilon(w)(v - u)\} \quad (5.13)$$

is a fine cover of N . This is how the full/fine arguments work. For, if not, then there would be some point x in E so that, for every $\varepsilon > 0$,

$$\beta_2 = \{([u, v], w) : w \in E, |F(v) - F(u)| < \varepsilon(v - u)\} \quad (5.14)$$

would have to be full at x . But that says exactly that $F'(x) = 0$. Write $N_i = \{w \in N : \varepsilon(w) > 1/i\}$ for each integer i and note that N is the union of the sequence of sets $\{N_i\}$.

Fix i . Let $\eta > 0$. Since F has zero variation on E we can find a full cover β_3 of E so that

$$\sum_{([u, v], w) \in \pi} |F(v) - F(u)| < \eta \quad (5.15)$$

whenever π is a subpartition, $\pi \subset \beta_3$. The intersection $\beta = \beta_1 \cap \beta_3$ is a fine cover of N .

For the set N_i and any subpartition $\pi \subset \beta[N_i]$ we compute, with some help from (5.13) and (5.15), that

$$\begin{aligned} \sum_{([u, v], w) \in \pi} (v - u) &< \sum_{([u, v], w) \in \pi} \varepsilon(w) |F(v) - F(u)| \\ &\leq i \sum_{([u, v], w) \in \pi} |F(v) - F(u)| < i\eta. \end{aligned}$$

This verifies that each set N_i is a fine null set and so, by the Mini-Vitali covering theorem, also a set of measure zero. Consequently N itself, as the union of a sequence of measure zero sets, is also a set of measure zero. This completes the proof. ■

5.6.2 Generalization of the zero derivative/variation

We wish to interpret this result in a much more general manner. Let h be any real-valued function that assigns values $h([u, v], w)$ to pairs $([u, v], w)$. We can define zero variation and zero derivative for h just as easily as we can for a

function $F : \mathbb{R} \rightarrow \mathbb{R}$.

If $h(I, x)$ is any function which assigns real values to interval-point pairs it will be convenient to have a notation for the following limits:

$$\limsup_{(I,x) \Rightarrow x} h(I, x) = \inf_{\delta > 0} (\sup \{h(I, x) : \lambda(I) < \delta, x \in I\})$$

and

$$\liminf_{(I,x) \Rightarrow x} h(I, x) = \sup_{\delta > 0} (\inf \{h(I, x) : \lambda(I) < \delta, x \in I\}).$$

These are just convenient expressions for the lower and upper limits of $h(I, x)$ as the interval I (always assumed to contain x) shrinks to the point x .

We say that h has a *zero derivative* at a point w if

$$\limsup_{(I,w) \Rightarrow w} \left| \frac{h(I, w)}{\lambda(I)} \right| = 0.$$

This is equivalent to requiring that

$$\lim_{\delta \rightarrow 0+} \sup \left\{ \left| \frac{h([u, v], w)}{v - u} \right| : u \leq w \leq v, 0 < v - u < \delta \right\} = 0.$$

We say too that h has *zero variation* on a set E if for every $\varepsilon > 0$ there is a full cover β of E so that

$$\sum_{([u,v],w) \in \pi} |h([u, v], w)| < \varepsilon$$

whenever π is a subpartition, $\pi \subset \beta$.

A repeat of the proofs just given, with minor changes, allows us to claim that Theorems 5.21 and 5.22 can be extended to these general versions:

Theorem 5.23 *If h has a zero derivative everywhere in a set E then h has zero variation on E .*

Theorem 5.24 *Zero variation for h on a set E implies h has a zero derivative almost everywhere in E .*

Exercise 628 Show that if

$$\limsup_{(I,x) \Rightarrow x} h(I,x) < t$$

at every point x of a set E then

$$\{(I,x) : x \in I, h(I,x) < t\}$$

is a full cover of E . □

Exercise 629 Show that if

$$\liminf_{(I,x) \Rightarrow x} h(I,x) < t$$

at every point x of a set E then

$$\{(I,x) : x \in I, h(I,x) < t\}$$

is a fine cover of E . □

5.6.3 Absolutely continuous functions

Our formulation of the notions of zero variation and full null set are immediately related by the fact that the function $F(x) = x$ has zero variation on a set N precisely when that set N is a set of measure zero. We see, then, that $F(x) = x$ has zero variation on all sets of measure zero. Most functions that we have encountered in the calculus also have this property. We shall see that all differentiable functions have this property. It plays a vital role in the theory; such functions are said to be absolutely continuous⁵.

We first introduced this notion in Chapter 4 and we repeat and review it here for convenience.

Definition 5.25 A function $F : (a,b) \rightarrow \mathbb{R}$ is said to be absolutely continuous on the open interval (a,b) if F has zero variation on every subset N of the interval that has measure zero.

Absolute continuity is stronger than continuity.

⁵**Note to the instructor:** this notion is strictly more general than the traditional notion (due to Vitali) of a function absolutely continuous on a compact interval $[a,b]$. In particular an absolutely continuous function in this sense need not have bounded variation.

Lemma 5.26 *If a function $F : (a, b) \rightarrow \mathbb{R}$ is absolutely continuous on the open interval (a, b) then F is continuous at each point of that interval.*

Proof. If F has zero variation on each measure zero subset of (a, b) then F has zero variation on any set $\{x_0\}$ containing a single point x_0 from that interval. If we translate what this would mean into ε, δ language we find that for every $\varepsilon > 0$ there must be a $\delta > 0$ so that

$$|F(v) - F(u)| < \varepsilon$$

if $v - u < \delta$ and $x_0 \in [u, v]$. But this is exactly the statement that F is continuous at the point x_0 . ■

The exercises show that most continuous functions we encounter in the calculus will be absolutely continuous. In fact the only continuous function we have seen so far that is not absolutely continuous is the Cantor function of Section 4.4.

5.6.4 Absolute continuity and derivatives

There is an intimate relationship between the differentiability properties of a function and its absolute continuity properties. The first such connection is easy to make. Our lemma shows that all differentiable functions are absolutely continuous.

Lemma 5.27 *Suppose that F is a real-valued function defined on an open set that contains the measure zero set N and that F is differentiable at every point of N . Then F has zero variation on N .*

Proof. For each natural number n let N_n be the collection of those points x in N at which $|F'(x)| < n$. We show that F has zero variation on each N_n . It follows then that F has zero variation on $N = \bigcup_{n=1}^{\infty} N_n$.

Let $\varepsilon > 0$. Since There must be a full cover β_1 of N so that

$$\sum_{([u,v],w) \in \pi} (v - u) < \varepsilon/n$$

whenever π is a subpartition, $\pi \subset \beta_1$. Define

$$\beta_2 = \{([u, v], w) : w \in E_n, |F(v) - F(u)| < n(v - u)\}.$$

This is evidently a full cover of N_n , because $|F'(w)| < n$ for each $w \in N_n$.

Consequently $\beta = \beta_1 \cap \beta_2$ is a full cover of N_n for which

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)| < \sum_{([u,v],w) \in \pi} n(v-u) < \varepsilon$$

whenever π is a subpartition, $\pi \subset \beta$. This proves that F has zero variation on N_n . Since N is the union of the sequence of set N_n this proves our assertion. ■

Exercises

Exercise 630 Show that the function $F(x) = x$ is absolutely continuous on every open interval. □

Exercise 631 Show that a linear combination of absolutely continuous functions is absolutely continuous. □

Exercise 632 Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is absolutely continuous on the interval (a, b) . Show that F must be pointwise continuous at every point of that interval. □

Exercise 633 Show that a Lipschitz function defined on an open interval is absolutely continuous there. □

Exercise 634 Give an example of an absolutely continuous function that is not Lipschitz. □

Exercise 635 Show that the Cantor function is not absolutely continuous on $(0, 1)$. □

Exercise 636 Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is differentiable at each point of the open interval (a, b) . Show that F is absolutely continuous on the interval (a, b) . Answer □

Exercise 637 Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is differentiable at each point of the open interval (a, b) with countably many exceptions but that F is pointwise continuous at those exceptional points. Show that F is absolutely continuous on the interval (a, b) . Answer □

Exercise 638 Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is differentiable at each point of the open interval (a, b) with the exception of a set $N \subset (a, b)$. Suppose further that N is a set of measure zero and that F has zero variation on N . Show that F is absolutely continuous on the interval (a, b) . $\square$

Exercise 639 Suppose that $F : (a, b) \rightarrow \mathbb{R}$ is absolutely continuous on the interval (a, b) . Then by definition F has zero variation on every subset of measure zero. Is it possible that F has zero variation on subsets that are not measure zero? $\square$

Exercise 640 A function F on an open interval I is said to have finite derived numbers on a set $E \subset I$ if, for each $x \in E$, there is a number M_x and one can choose $\delta > 0$ so that

$$\left| \frac{F(x+h) - F(x)}{h} \right| \leq M_x$$

whenever $x+h \in I$ and $|h| < \delta$. Show that F is absolutely continuous on E if F has finite derived numbers there. [cf. Exercise 170.] $\square$

5.7 Lebesgue differentiation theorem

Our second application of the Mini-Vitali theorem is to prove a famous and useful theorem of Lebesgue asserting that functions of bounded variation are almost everywhere differentiable. We shall need this in our study of the Lebesgue integral in Chapter 7.

Theorem 5.28 Let $F : [a, b] \rightarrow \mathbb{R}$ be a function of bounded variation. Then F is differentiable at almost every point in (a, b) .

Corollary 5.29 Let $F : [a, b] \rightarrow \mathbb{R}$ be a monotonic function. Then F is differentiable at almost every point in (a, b) .

The proof of the theorem will require an introduction, first, to the upper and lower derivatives and then a simple geometric lemma that allows us to use a fine covering argument to show that the set of points where $F'(x)$ does not exist is measure zero.

5.7.1 Upper and lower derivates

The proof uses the *upper and lower derivates*. To analyze how a derivative $F'(x)$ may fail to exist we split that failure into two pieces, an upper and a lower, defined as

$$\overline{DF}(x) = \inf_{\delta > 0} \sup \left\{ \frac{F(v) - F(u)}{v - u} : x \in [u, v], 0 < v - u < \delta \right\}$$

and

$$\underline{DF}(x) = \sup_{\delta > 0} \inf \left\{ \frac{F(v) - F(u)}{v - u} : x \in [u, v], 0 < v - u < \delta \right\}$$

We will prove that, for almost every point x in (a, b) ,

$$\overline{DF}(x) > -\infty, \quad \underline{DF}(x) < \infty,$$

and

$$\overline{DF}(x) = \underline{DF}(x).$$

From these three assertions it follows that F has a finite derivative $F'(x)$ at almost every point x in (a, b) .

The proof will depend on a fine covering argument. For that we need to recognize the following connection between derivates and covers. The proof is trivial; it is only a matter of interpreting the statements.

Lemma 5.30 *Let $F : [a, b] \rightarrow \mathbb{R}$, $\alpha \in \mathbb{R}$, and let*

$$\beta = \left\{ ([u, v], w) : \frac{F(v) - F(u)}{v - u} > \alpha, w \in [u, v] \subset [a, b] \right\}.$$

Then, β is a full cover of the set

$$E_1 = \{x \in (a, b) : \underline{DF}(x) > \alpha\}$$

and a fine cover of the larger set

$$E_2 = \{x \in (a, b) : \overline{DF}(x) > \alpha\}.$$

5.7.2 Geometrical lemmas

The proof employs an elementary geometric lemma that Donald Austin⁶ used in 1965 to give a simple proof of this theorem. Our proof of the differentiation theorem is essentially his, but written in different language. See also the version of Michael Botsko⁷.

Lemma 5.31 (Austin's lemma) *Let $G : [a, b] \rightarrow \mathbb{R}$, $\alpha > 0$ and suppose that $G(a) \leq G(b)$. Let*

$$\beta = \left\{ ([u, v], w) : \frac{G(v) - G(u)}{v - u} < -\alpha, w \in [u, v] \subset [a, b] \right\}.$$

Then, for any nonempty subpartition $\pi \subset \beta$,

$$\alpha \left(\sum_{([u, v], w) \in \pi} (v - u) \right) < V(G, [a, b]) - |G(b) - G(a)|.$$

Proof. To prove the lemma, let π_1 be a partition of $[a, b]$ that contains the subpartition π . Just write

$$\begin{aligned} |G(b) - G(a)| &= G(b) - G(a) = \sum_{([u, v], w) \in \pi_1} [G(v) - G(u)] \\ &= \sum_{([u, v], w) \in \pi} [G(v) - G(u)] + \sum_{([u, v], w) \in \pi_1 \setminus \pi} [G(v) - G(u)] \\ &< -\alpha \left(\sum_{([u, v], w) \in \pi} [v - u] \right) + V(G, [a, b]). \end{aligned}$$

The statement of the lemma follows. ■

As a corollary we can replace F with $-F$ to obtain a similar statement.

⁶D. Austin, *A geometric proof of the Lebesgue differentiation theorem*. Proc. Amer. Math. Soc. 16 (1965) 220–221.

⁷M. W. Botsko, *An elementary proof of Lebesgue's differentiation theorem*. Amer. Math. Monthly 110 (2003), no. 9, 834–838.

Corollary 5.32 Let $G : [a, b] \rightarrow \mathbb{R}$, $\alpha > 0$ and suppose that $G(b) \leq G(a)$. Let

$$\beta = \left\{ ([u, v], w) : \frac{G(v) - G(u)}{v - u} > \alpha, w \in [u, v] \subset [a, b] \right\}.$$

Then, for any nonempty subpartition $\pi \subset \beta$,

$$\alpha \left(\sum_{([u, v], w) \in \pi} (v - u) \right) < V(G, [a, b]) - |G(b) - G(a)|.$$

5.7.3 Proof of the Lebesgue differentiation theorem

We now prove the theorem. The first step in the proof is to show that at almost every point t in (a, b) ,

$$\underline{DF}(t) = \overline{DF}(t).$$

If this is not true then there must exist a pair of rational numbers r and s for which the set

$$E_{rs} = \{t \in (a, b) : \underline{DF}(t) < r < s < \overline{DF}(t)\}$$

is not a set of measure zero. This is because the union of the countable collection of sets E_{rs} contains all points t for which $\underline{DF}(t) \neq \overline{DF}(t)$.

Let us show that each such set E_{rs} is fine null. By the Mini-Vitali theorem we then know that E_{rs} is a set of measure zero. Write $\alpha = (s - r)/2$, $B = (r + s)/2$, $G(t) = F(t) - Bt$. Note that

$$E_{rs} = \{t \in (a, b) : \underline{DG}(t) < -\alpha < 0 < \alpha < \overline{DG}(t)\}.$$

Since F has bounded variation on $[a, b]$, so too does the function G . In fact

$$V(G, [a, b]) \leq V(F, [a, b]) + B(b - a).$$

Let $\varepsilon > 0$ and select points

$$a = s_0 < s_1 < \cdots < s_{n-1} < s_n = b$$

so that

$$\sum_{i=1}^n |G(s_i) - G(s_{i-1})| > V(G, [a, b]) - \alpha\varepsilon.$$

Let $E'_{rs} = E_{rs} \setminus \{s_1, s_2, \dots, s_{n-1}\}$. Let us call an interval $[s_{i-1}, s_i]$ *black* if $G(s_i) - G(s_{i-1}) \geq 0$ and call it *red* if

$$G(s_i) - G(s_{i-1}) < 0.$$

For each $i = 1, 2, 3, \dots, n$ we define a covering relation β_i as follows. If $[s_{i-1}, s_i]$ is a black interval then

$$\beta_i = \left\{ ([u, v], w) : \frac{G(v) - G(u)}{v - u} < -\alpha, w \in [u, v] \subset [s_{i-1}, s_i] \right\}.$$

If, instead, $[s_{i-1}, s_i]$ is a red interval then

$$\beta_i = \left\{ ([u, v], w) : \frac{G(v) - G(u)}{v - u} > \alpha, w \in [u, v] \subset [s_{i-1}, s_i] \right\}.$$

Let $\beta = \bigcup_{i=1}^n \beta_i$. Because of Lemma 5.30 we see that this collection β is a fine cover of E'_{rs} .

Let π be any nonempty subpartition contained in β . Write

$$\pi_i = \pi \cap \beta_i.$$

By Lemma 5.31 applied to the black intervals and Corollary 5.31 applied to the red intervals we obtain that

$$\alpha \left(\sum_{([u, v], w) \in \pi_i} (v - u) \right) < V(G, [s_{i-1}, s_i]) - |G(s_i) - G(s_{i-1})|.$$

Consequently

$$\begin{aligned} \alpha \left(\sum_{([u, v], w) \in \pi} (v - u) \right) &= \alpha \left(\sum_{i=1}^n \sum_{([u, v], w) \in \pi_i} (v - u) \right) \\ &\leq \sum_{i=1}^n V(G, [s_{i-1}, s_i]) - \sum_{i=1}^n |G(s_i) - G(s_{i-1})| \\ &\leq V(G, [a, b]) - [V(G, [a, b]) - \alpha \varepsilon] = \alpha \varepsilon. \end{aligned}$$

We have proved that β is a fine cover of E'_{rs} with the property that

$$\sum_{([u, v], w) \in \pi} (v - u) < \varepsilon$$

for every subpartition $\pi \subset \beta$. It follows that E'_{rs} is fine null, and hence a set of measure zero. So too then is E_{rs} since the two sets differ by only a finite number of points.

We know now that the function F has a derivative, finite or infinite, almost everywhere in (a, b) . We wish to exclude

the possibility of the infinite derivative, except on a set of measure zero.

Let

$$E_{\infty} = \{t \in (a, b) : \underline{D}F(t) = \infty\}.$$

Choose any B so that $F(b) - F(a) \leq B(b - a)$ and set $G(t) = F(t) - Bt$. Note that $G(b) \leq G(a)$ which will allow us to apply Corollary 5.32.

Let $\varepsilon > 0$ and choose a positive number α large enough so that

$$V(G, [a, b]) - |G(b) - G(a)| < \alpha\varepsilon.$$

Define

$$\beta = \left\{ ([u, v], w) : \frac{G(v) - G(u)}{v - u} > \alpha, [u, v] \subset [a, b] \right\}.$$

This is a fine cover of E_{∞} . Let π be any subpartition $\pi \subset \beta$. By our corollary then

$$\alpha \sum_{([u, v], w) \in \pi} (v - u) < V(G, [a, b]) - |G(b) - G(a)| < \alpha\varepsilon.$$

We have proved that β is a fine cover of E_{∞} with the property that

$$\sum_{([u, v], w) \in \pi_i} (v - u) < \varepsilon$$

for every subpartition $\pi \subset \beta$. It follows that E_{∞} is fine null, and hence a set of measure zero. The same arguments will handle the set

$$E_{-\infty} = \{t \in (a, b) : \overline{D}F(t) = -\infty\}.$$

Exercise 641 *The formula*

$$\frac{d}{dx} \sum_{n=1}^{\infty} F_n(x) = \sum_{n=1}^{\infty} \frac{d}{dx} F_n(x)$$

is not generally valid without assumptions about uniform convergence (see Chapter 3). Fubini's differentiation theorem says that, with some assumptions on the nature of the functions F_n , we can have this differentiation formula, not everywhere, but almost everywhere. Prove this as an application of the Lebesgue differentiation theorem:

Theorem 5.33 (Fubini) Let $\{F_n\}$ be a sequence of monotonic, nondecreasing functions on the interval $[a, b]$ and suppose that $F(x) = \sum_{n=1}^{\infty} F_n(x)$ is absolutely convergent for all $a \leq x \leq b$. Then, for almost every x in (a, b) ,

$$F'(x) = \sum_{n=1}^{\infty} F'_n(x).$$

Answer $\square$

Chapter 6

The Integral

This chapter studies the natural integral on the real line. We started our study of the definite integral in Chapter 3 with this limited definition.

Definition 6.1 (Naive calculus integral) *By the statement*

$$\int_a^b F'(x) dx = F(b) - F(a)$$

we mean that $F : [a, b] \rightarrow \mathbb{R}$ is uniformly continuous and differentiable mostly everywhere.

In Chapter 4 we replaced this with the correct version. We repeat that definition now and we shall review some of the definitions and properties in this chapter.

Definition 6.2 (The calculus integral) *By the statement*

$$\int_a^b F'(x) dx = F(b) - F(a)$$

we mean that $F : [a, b] \rightarrow \mathbb{R}$ is uniformly continuous and differentiable almost everywhere, and is absolutely continuous on (a, b) .

To see that the second definition is a generalization of the first requires us to recall that a continuous function that is differentiable mostly everywhere on an open interval (a, b) is absolutely continuous there (Exercise 503).

The naive version is strong enough for a wide variety of applications and is certainly accessible to calculus students since it depends only on an understanding of the derivative, continuity, and the mean-value theorem. One discovers that the definition is “naive” only when one begins to work on harder problems that arise in working with sequences and series of integrable functions.

One advantage we shall have in our studies is that the naive calculus integral is a very good teaching tool in preparing us to study the correct version of integration theory on the real line. It leads by mostly easy proofs and natural arguments to the full theory without too much detailed development. The Lebesgue integral will appear in Chapter 7 but our presentation should prove a little easier than the usual introductions to that integral. If one starts with Lebesgue’s definition it is necessary first to prove all properties of the integral using that definition. Since we already have an integration theory developed, we need only check that Lebesgue’s methods can be used to construct the value of the integral.

6.1 The integral and integrable functions

Here is the formal definition of the integral. We shall not necessarily always call it a “calculus integral” although that is the intent. Most authors would call it a “descriptive definition” of the integral. It describes the situation of integrability without offering any methods for the computation of the integral. There is an illusion of constructibility since the value is given as $F(b) - F(a)$ but, in fact, the definition offers us no method for ever finding such a function F .

Definition 6.3 (Definite integral) Let $f : (a, b) \rightarrow \mathbb{R}$ be a function defined at all points of the open interval (a, b) with the possible exception of a set of measure zero. Then f is said to be integrable on the compact interval $[a, b]$ provided there is a function $F : [a, b] \rightarrow \mathbb{R}$ so that

1. F is uniformly continuous on $[a, b]$.
2. F is absolutely continuous on (a, b) .
3. $F'(x) = f(x)$ at all points x of (a, b) with the possible exception of a set of measure zero.

In that case we define $\int_a^b f(x) dx = F(b) - F(a)$.

Sometimes it is more convenient to state the conditions for the integral with direct attention to the set of exceptional points where the derivative $F'(x) = f(x)$ may fail. Conditions 1, 2, and 3 of the definition can be replaced by requiring instead that

1. F is uniformly continuous on $[a, b]$.
2. There is a set N of measure zero.
3. $F'(x) = f(x)$ at all points x of (a, b) with the possible exception of points in N .
4. F has zero variation on N .

Exercise 504 can be used to show that these four statements are equivalent to the three statements in the definition.

6.1.1 Infinite integrals

Exactly the same definition for the infinite integrals

$$\int_{-\infty}^{\infty} f(x) dx, \quad \int_a^{\infty} f(x) dx, \quad \text{and} \quad \int_{-\infty}^b f(x) dx$$

can be given as for the integral over a compact interval.

Definition 6.4 Let f be a function defined at every point of $(-\infty, \infty)$ with the possible exception of a set of measure zero. Then f is said to be integrable on $(-\infty, \infty)$ provided there is a function $F : (-\infty, \infty) \rightarrow \mathbb{R}$ so that

1. F is absolutely continuous on $(-\infty, \infty)$.
2. $F'(x) = f(x)$ at all points x with the possible exception of a set of measure zero.
3. Both limits $F(\infty) = \lim_{x \rightarrow \infty} F(x)$ and $F(-\infty) = \lim_{x \rightarrow -\infty} F(x)$ exist.

In that case the number $\int_{-\infty}^{\infty} f(x) dx = F(\infty) - F(-\infty)$, is called the definite integral of f on the interval $(-\infty, \infty)$.

Similar assertions define

$$\int_{-\infty}^b f(x) dx = F(b) - F(-\infty)$$

and

$$\int_a^{\infty} f(x) dx = F(\infty) - F(a).$$

In analogy with the terminology of an infinite series $\sum_{k=1}^{\infty} a_k$ we often say that the integral $\int_a^{\infty} f(x) dx$ converges when the integral exists. That suggests language asserting that the integral converges absolutely if both integrals

$$\int_a^{\infty} f(x) dx \quad \text{and} \quad \int_a^{\infty} |f(x)| dx$$

exist.

6.1.2 Approximation by Riemann sums

We have seen in Chapter 3 that all naive calculus integrals can be approximated by Riemann sums, *pointwise approximated* that is. The same theorem is true for the advanced integration theory. The proof is elementary if detailed.

Theorem 6.5 (Henstock's criterion) Suppose that f is an integrable function defined at every point of a compact interval $[a, b]$. Then for every $\varepsilon > 0$ there is a full cover β of $[a, b]$ so that

$$\sum_{[u,v],w \in \pi} \left| \int_u^v f(x) dx - f(w)(v-u) \right| < \varepsilon$$

and

$$\left| \int_a^b f(x) dx - \sum_{[u,v],w \in \pi} f(w)(v-u) \right| < \varepsilon$$

whenever π is a partition of the interval $[a, b]$ chosen from β .

Note that the function here is defined at every point of the interval $[a, b]$. We do not usually insist on this, permitting instead integrable functions to be defined only almost everywhere. The way to make this theorem accessible in general is assign arbitrary values to the function at points where it is undefined. This does not alter integrability or change the integral in any way. A frequent convention, given a function f defined almost everywhere on an interval (a, b) , is to work instead with the function g where we take $g(x) = f(x)$ when that exists and $g(x) = 0$ otherwise.

Proof. Let F be an indefinite integral for f on the interval $[a, b]$. We can set $F(x) = F(a)$ for $x < a$, $F(x) = F(b)$ for $x > b$ and set $f(x) = 0$ outside of (a, b) . We suppose that $F'(x) = f(x)$ for all x excepting a set N of measure zero. Write, for every integer $j = 1, 2, 3, \dots$,

$$N_j = \{x \in N : j-1 \leq |f(x)| < j\}$$

and note that the set N is the union of the sequence of sets $\{N_j\}$. Each N_j is a measure zero set and so there is a full cover β_j of N_j so that

$$\left| \sum_{([u,v],w) \in \pi} (v-u) \right| < \varepsilon 2^{-j-2} j^{-1}$$

whenever π is a subpartition chosen from β_j . Note that

$$\sum_{([u,v],w) \in \pi} |f(w)|(v-u) < \varepsilon 2^{-j-1}$$

whenever π is a subpartition chosen from $\beta_j[N_j]$. Since N is measure zero and F is absolutely continuous we know that

F has zero variation on N . Consequently there is full cover β' of N so that

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)| < \varepsilon/4$$

whenever π is a subpartition chosen from β' .

Finally $F'(x) = f(x)$ for all $x \notin N$. Thus

$$\beta'' = \left\{ ([u,v],w) : |F(v) - F(u) - f(w)(v-u)| \leq \varepsilon \frac{v-u}{2(b-a)} \right\}$$

is a full cover of $\mathbb{R} \setminus N$.

Now we can construct our full cover β needed in the statement of the theorem. Let

$$\beta = \beta'' \cup \left(\beta'[N] \cap \bigcup_{j=1}^{\infty} \beta_j[N_j] \right).$$

Let π be a partition of the interval $[a, b]$ chosen from β and estimate

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u) - f(w)(v-u)|.$$

If w is a point where $F'(w) = f(w)$ then

$$|F(v) - F(u) - f(w)(v-u)| \leq \varepsilon \frac{v-u}{2(b-a)}$$

and the contribution to the sum of such terms is evidently smaller than $\varepsilon/2$.

If w is not a point where $F'(w) = f(w)$ then we can treat the sum of such terms by estimating using the larger value

$$|F(v) - F(u) - f(w)(v-u)| \leq |F(v) - F(u)| + |f(w)|(v-u).$$

The sum of the terms

$$|F(v) - F(u)|$$

where $w \in N$ cannot exceed $\varepsilon/4$. The sum of the terms

$$|f(w)|(v-u)$$

where $w \in N_j$ cannot exceed $\varepsilon 2^{-j-1}$. Putting these together shows that

$$\sum_{[u,v],w \in \pi} \left| \int_u^v f(x) dx - f(w)(v-u) \right| \leq \sum_{[u,v],w \in \pi} |F(v) - F(u) - f(w)(v-u)| < \varepsilon$$

as required. The final inequality of the theorem follows directly from this inequality since

$$\left| \int_a^b f(x) dx - \sum_{[u,v],w \in \pi} f(w)(v-u) \right| \leq \sum_{[u,v],w \in \pi} \left| \int_u^v f(x) dx - f(w)(v-u) \right|$$

follows immediately from the triangle inequality. ■

6.2 The Henstock-Kurzweil characterization of the integral

We turn to the converse direction suggested by the preceding section. Since every calculus integral can be captured by the Riemann sums approach we ask whether the Riemann sums approach captures the calculus integral itself. The answer is yes and we devote all of this section to exploring this relation. At the end we will have two tools for working with our integration theory: the methods of the calculus integral and the methods of Henstock and Kurzweil.

6.2.1 Definition of Henstock and Kurzweil

So far we have the Henstock criterion as a necessary condition for integrability of a function. Our goal now is to show that it is also a sufficient condition for integrability. That then gives us two very different ways of completely describing the integral. Our first way is to state the integral as a calculus integral that merely interprets conditions under which

$$\int_a^b F'(x) dx = F(b) - F(a)$$

is valid. The second describes the integral as a limit of Riemann sums, similar to the methods that Riemann originally used to describe his integral.

Definition 6.6 (Henstock-Kurzweil) Suppose that f is a function defined at every point of a compact interval $[a, b]$. Then f is integrable in the Henstock-Kurzweil sense if there is a real number I such that for every $\varepsilon > 0$ there is a full cover β of $[a, b]$ so that

$$\left| I - \sum_{[u,v],w \in \pi} f(w)(v-u) \right| < \varepsilon$$

whenever π is a partition of the interval $[a, b]$ chosen from β . In that case the number I is set equal to

$$I = \int_a^b f(x) dx$$

and this integral is said to be interpreted in the Henstock-Kurzweil sense.

6.2.2 Upper and lower integrals

The Henstock-Kurzweil integral can be studied by means of an upper and a lower integral. This is a useful way to develop the theory and so we can leave Definition 6.6 behind us for a moment and start the theory of this integral in this way. This notion of using upper and lower integrals goes back at least to 1875 and is due to Jean-Gaston Darboux (1842–1917).

Definition 6.7 For a function $f : [a, b] \rightarrow \mathbb{R}$ we define an upper integral by

$$\overline{\int_a^b} f(x) dx = \inf_{\beta} \sup_{\pi \subset \beta} \left\{ \sum_{([u,v],w) \in \pi} f(w)(v-u) \right\}$$

where the supremum is taken over all partitions π of $[a, b]$ contained in β , and the infimum over all full covers β of the interval $[a, b]$.

Note that the first step is to estimate the largest possible value for the Riemann sums for partitions π of $[a, b]$ contained in β , and the second step is to refine this by shrinking to smaller and smaller full covers β .

Similarly we define a lower integral as

$$\int_a^b f(x) dx = \sup_{\beta} \inf_{\pi \in \beta} \left\{ \sum_{([u,v],w) \in \pi} f(w)(v-u) \right\}$$

where, again, π is a partition of $[a, b]$ and β is a full cover.

Exercises

Exercise 642 Check that, for any full cover β ,

$$-\infty < \sup_{\pi \in \beta} \sum_{([u,v],w) \in \pi} f(w)(v-u).$$

□

Exercise 643 Check that

$$\int_a^b f(x) dx = - \overline{\int_a^b [-f(x)] dx}.$$

□

Exercise 644 Let $f : [a, b] \rightarrow \mathbb{R}$. Show that

$$\int_a^b f(x) dx \leq \overline{\int_a^b f(x) dx}.$$

Answer □

Exercise 645 Show that a function f can be altered at a finite number of points without altering the values of the upper and lower integrals.

□

Exercise 646 Show that a function f can be altered at a countable number of points without altering the values of the upper and lower integrals.

□

Exercise 647 Let $f : [a, c] \rightarrow \mathbb{R}$ and suppose that $a < b < c$. Show that

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx,$$

assuming the sum makes sense.

Answer $\square$

Exercise 648 Let $f, g : \mathbb{R} \rightarrow \mathbb{R}$. What rule should hold for the upper and lower integrals

$$\int_a^b [f(x) + g(x)] dx \quad \text{and} \quad \int_a^b [f(x) + g(x)] dx?$$

$\square$

Exercise 649 Define a partition π to be endpointed if only elements of the form $([u, w], w)$ or $([w, v], w)$ appear and there is no element $([u, v], w) \in \pi$ for which $u < w < v$. Show that a restriction in the definition of integrals to use endpointed partitions only would not change the theory at all.

Answer $\square$

6.2.3 The integral and integrable functions

If the upper and lower Henstock-Kurzweil integrals are identical we write the common value as

$$\int_a^b f(x) dx = \int_a^b f(x) dx = \int_a^b f(x) dx$$

allowing finite or infinite values. We say in this case that the integral is *determined*. When the integral is not determined then (by Exercise 642)

$$\int_a^b f(x) dx < \int_a^b f(x) dx$$

and there is no integral.

If the integral is determined and this value is also finite then f is *Henstock-Kurzweil integrable* and

$$\int_a^b f(x) dx$$

is called the *Henstock-Kurzweil integral*, now assuming a finite value. Our first goal will be to check that this account is equivalent to Definition 6.6.

Exercises

Exercise 650 Let $f : [a, b] \rightarrow \mathbb{R}$ show that a sufficient condition for f to be Henstock-Kurzweil integrable on $[a, b]$ with $c = \int_a^b f(x) dx$ is that for every $\varepsilon > 0$ there is a full cover so that

$$\left| c - \sum_{([u,v],w) \in \pi} f(w)(v-u) \right| < \varepsilon$$

for every partition π of $[a, b]$ contained in β .

Answer $\square$

Exercise 651 Let $f : [a, b] \rightarrow \mathbb{R}$ be a Henstock-Kurzweil integrable function and let π be any partition of $[a, b]$. Show that

$$\left| \int_a^b f(x) dx - \sum_{([u,v],w) \in \pi} f(w)(v-u) \right| \leq \sum_{([u,v],w) \in \pi} \omega f([u, v]) \lambda([u, v]).$$

Here $\omega f(I)$ denotes the oscillation of the function f on the interval I , defined as

$$\sup_{s,t \in I} |f(s) - f(t)|.$$

$\square$

Exercise 652 Show that a Henstock-Kurzweil integrable function f can be altered at a finite number of points without altering the value of the integral. $\square$

Exercise 653 Show that a Henstock-Kurzweil integrable function f can be altered at a countable number of points without altering the value of the integrals. $\square$

Exercise 654 Define a function to be uniformly integrable [i.e., Riemann integrable] if in the definition one uses the uniformly full covers from Section 5.1.6, rather than the more general full covers. Show that a function that is integrable in this narrow sense must be bounded. $\square$

Exercise 655 Continuing the study of the Riemann integral begun in the preceding exercise, show that a function f that is uniformly integrable on an interval $[a, b]$ must satisfy the following restrictive property: for every $\varepsilon > 0$ there must

exist a partition π for which

$$\sum_{([u,v],w) \in \pi} \omega f([u,v])(v-u) < \varepsilon.$$

□

Exercise 656 Continuing the preceding two exercises (if you have the patience to work this hard on the Riemann integral), show that a function f is uniformly integrable on an interval $[a, b]$ if and only if it is bounded and satisfies the following property: for every $\varepsilon > 0$ there must exist a partition π for which

$$\sum_{([u,v],w) \in \pi} \omega f([u,v])(v-u) < \varepsilon.$$

□

6.2.4 First Cauchy criterion

Our first criterion for integrability returns us to Definition 6.6 and shows that the upper/lower integral approach is equivalent to the original Henstock-Kurzweil definition.

Theorem 6.8 A necessary and sufficient condition in order for a function $f : [a, b] \rightarrow \mathbb{R}$ to be Henstock-Kurzweil integrable on a compact interval $[a, b]$ is that there is a number c so that for all $\varepsilon > 0$ a full cover β of $[a, b]$ can be found so that

$$\left| \sum_{([u,v],w) \in \pi} f(w)(v-u) - c \right| < \varepsilon$$

for all partitions π of $[a, b]$ contained in β .

Proof. In Exercise 650 we checked that this condition is sufficient. On the other hand, if we know that f is integrable with $c = \int_a^b f(x) dx$ then, using the definition of the upper integral, for any $\varepsilon > 0$ we choose a full cover β_1 so that

$$\sum_{([u,v],w) \in \pi} f(w)(v-u) < c + \varepsilon$$

for all partitions π of $[a, b]$ contained in β_1 . Similarly, using the definition of the lower integral, we choose a full cover

β_2 so that

$$\sum_{([u,v],w) \in \pi} f(w)(v-u) > c - \varepsilon$$

for all partitions π of $[a, b]$ contained in β_2 . Take $\beta = \beta_1 \cap \beta_2$. This is a full cover with the property stated. ■

6.2.5 Second Cauchy criterion

Theorem 6.9 *A necessary and sufficient condition in order for a function $f : [a, b] \rightarrow \mathbb{R}$ to be Henstock-Kurzweil integrable on a compact interval $[a, b]$ is that, for all $\varepsilon > 0$, a full cover β of $[a, b]$ can be found so that*

$$\left| \sum_{(I,w) \in \pi} \sum_{(I',w') \in \pi'} [f(w) - f(w')] \lambda(I \cap I') \right| < \varepsilon \quad (6.1)$$

for all partitions π, π' of $[a, b]$ contained in β .

Proof. Start by checking that when π and π' are both partitions of the same interval $[a, b]$ then, for any subinterval I of $[a, b]$

$$\lambda(I) = \sum_{(I',w') \in \pi'} \lambda(I \cap I')$$

from which it is easy to see that

$$\sum_{(I,w) \in \pi} f(w) \lambda(I) = \sum_{(I,w) \in \pi} \sum_{(I',w') \in \pi'} f(w) \lambda(I \cap I').$$

This allows the difference that would normally appear in a Cauchy type criterion

$$\left| \sum_{(I,w) \in \pi} f(w) \lambda(I) - \sum_{(I',w') \in \pi'} f(w') \lambda(I') \right|$$

to assume the simple form given in (6.1). In particular that statement can be rewritten as

$$\left| \sum_{(I,w) \in \pi} f(w) \lambda(I) - \sum_{(I',w') \in \pi'} f(w') \lambda(I') \right| < \varepsilon. \quad (6.2)$$

The condition is necessary. For if f is integrable then the first Cauchy criterion supplies a full cover β so that

$$\left| \sum_{(I,w) \in \pi} f(w)\lambda(I) - c \right| < \varepsilon/2$$

for all partitions π of $[a, b]$ contained in β . Any two Riemann sums would both be this close to c and hence within ε of each other.

Suppose the condition holds. We can see from (6.2) that the upper and lower integrals must be finite. We wish to show that they are equal.

Using the definition of the upper integral, there is at least one partition π of $[a, b]$ contained in β with

$$\sum_{(I,w) \in \pi} f(w)\lambda(I) > \overline{\int_a^b} f(x) dx - \varepsilon$$

Using the definition of the lower integral, there is at least one partition π' of $[a, b]$ contained in β with

$$\sum_{(I,w) \in \pi'} f(w)\lambda(I) < \underline{\int_a^b} f(x) dx + \varepsilon.$$

Together with (6.2) these show that

$$\overline{\int_a^b} f(x) dx - \underline{\int_a^b} f(x) dx < 2\varepsilon.$$

Since ε is an arbitrary positive number the upper and lower integrals are equal. ■

Exercise 657 (McShane's criterion) A function $f : [a, b] \rightarrow \mathbb{R}$ is said to satisfy McShane's criterion on $[a, b]$ provided that for all $\varepsilon > 0$ a full cover β can be found so that

$$\sum_{(I,w) \in \pi} \sum_{(I',w') \in \pi'} |f(w) - f(w')| \lambda(I \cap I') < \varepsilon$$

for all partitions π, π' of $[a, b]$ contained in β . Show that if a function satisfies this criterion then both f and $|f|$ are integrable on $[a, b]$.

Note: the converse is proved in Chapter 7.

Answer □

6.2.6 Proof of equivalence

Our goal now is to prove that the Henstock-Kurzweil integral is exactly the same as the general calculus integral. One direction is simple and we have already stated it in Henstock's criterion (Theorem 715).

Here is the outline of the proof in this section. If the goal is only to effect a proof then the steps are not of much importance in themselves. They are interesting, however, for a different reason: if one wished to start with Definition 6.6 and use the Henstock-Kurzweil integral without first defining the calculus integral, then these are the first steps in developing the theory of that integral. In particular this section outlines the proof of the fundamental theorem of the calculus for the Henstock-Kurzweil integral.

Step 1 Henstock-Kurzweil integrability on subintervals.

Step 2 The Henstock criterion for the Henstock-Kurzweil indefinite integral.

Step 3 The Henstock criterion implies uniform continuity and absolute continuity of the indefinite Henstock-Kurzweil integral.

Step 4 The Henstock criterion implies the almost everywhere differentiability of the indefinite Henstock-Kurzweil integral.

Lemma 6.10 (HK integrability on subintervals) *If $f : [a, b] \rightarrow \mathbb{R}$ is Henstock-Kurzweil integrable then it is also integrable on any compact subinterval of $[a, b]$.*

Proof. Let $\varepsilon > 0$. Suppose that f is Henstock-Kurzweil integrable on $[a, b]$ and $[c, d]$ is a compact subinterval. Take any full cover β so that the second Cauchy criterion is satisfied for β .

Observe that for every pair of partitions π_1 , and $\pi_2 \subset \beta$ of the subinterval $[c, d]$, there is a subpartition π from β so that $\pi_1 \cup \pi$ and $\pi_2 \cup \pi$ are partitions of the full interval $[a, b]$. In particular then

$$\left| \sum_{(I,w) \in \pi_1} f(w)\lambda(I) - \sum_{(I,w) \in \pi_2} f(w)\lambda(I) \right| =$$

$$\left| \sum_{(I,w) \in \pi \cup \pi_1} f(w)\lambda(I) - \sum_{(I,w) \in \pi \cup \pi_2} f(w)\lambda(I) \right| < \varepsilon$$

The integrability of f on $[c, d]$ follows now from the second Cauchy criterion. ■

Lemma 6.11 (The indefinite HK integral) *If $f : [a, b] \rightarrow \mathbb{R}$ is Henstock-Kurzweil integrable then there is a function $F : [a, b] \rightarrow \mathbb{R}$, called an indefinite integral for f , so that*

$$\int_c^d f(x) dx = F(d) - F(c)$$

for every compact subinterval $[c, d]$ of $[a, b]$.

Proof. Lemma 6.10 supplies the existence of the integral on the subintervals. Then the function

$$F(t) = \int_a^t f(x) dx \quad (a \leq t \leq b)$$

will have this property.

To see this first check that if $a < c < d \leq b$ then

$$\int_a^c f(x) dx + \int_c^d f(x) dx = \int_a^d f(x) dx. \quad (6.3)$$

Consequently

$$\int_c^d f(x) dx = \int_a^d f(x) dx - \int_a^c f(x) dx = F(d) - F(c)$$

as we require. Thus the remainder of the proof is devoted to proving the identity (6.3). We will leave as an exercise to the reader to attempt this using the first Cauchy criterion. This also follows from Exercise 647. ■

Lemma 6.12 (Henstock's criterion for the HK integral) *A necessary and sufficient condition for a function $f : [a, b] \rightarrow \mathbb{R}$ to be Henstock-Kurzweil integrable on a compact interval $[a, b]$ and for F to be its indefinite integral is that for every $\varepsilon > 0$ there exists a full cover β of $[a, b]$ such that*

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u) - f(w)(v-u)| < \varepsilon, \quad (6.4)$$

for every subpartition π of $[a, b]$ contained in β .

Proof. Suppose that this criterion holds. Then (6.4) immediately shows that

$$\begin{aligned} & \left| F(b) - F(a) - \sum_{([u,v],w) \in \pi} f(w)(v-u) \right| \\ &= \left| \sum_{([u,v],w) \in \pi} [F(v) - F(u) - f(w)(v-u)] \right| \\ &\leq \sum_{([u,v],w) \in \pi} |F(v) - F(u) - f(w)(v-u)| < \varepsilon. \end{aligned}$$

It follows that $F(b) - F(a) = \int_a^b f(x) dx$ by the first Cauchy criteria. The same argument will work on any subinterval to check that F is an indefinite integral for f .

Conversely let us suppose that F is an indefinite integral for f on $[a, b]$ and $\varepsilon > 0$. By the Cauchy criterion there is a full cover β such that

$$\left| F(b) - F(a) - \sum_{([u,v],w) \in \pi} f(w)(v-u) \right| < \varepsilon/4 \quad (6.5)$$

for every partition π of $[a, b]$ contained in β and it will be our goal to establish (6.4) from this.

Fix π and let $\pi' \subset \pi$ be any nonempty subset. Since β is full and contains partitions of any compact interval, we will find a useful way to supplement the subpartition π' so as to form a useful partition of $[a, b]$: we write

$$\pi \setminus \pi' = \{([u_1, v_1], w_1), ([u_2, v_2], w_2), \dots, ([u_k, v_k], w_k)\}.$$

Our hypothesis requires F to be an indefinite integral for f on each $[u_i, v_i]$ ($i = 1, 2, \dots, k$) and so for each $i = 1, 2, \dots, k$ we are able to select a partition $\pi_i \subset \beta$ of the interval $[u_i, v_i]$ in such a way that

$$\left| F(v_i) - F(u_i) - \sum_{([u,v],w) \in \pi_i} f(w)(v-u) \right| < \varepsilon/(4k). \quad (6.6)$$

Thus if we augment π' to form

$$\pi'' = \pi \cup \pi_1 \cup \pi_2 \cup \dots \cup \pi_k$$

we obtain a partition of $[a, b]$ contained in β and thus also satisfying an inequality of the form (6.5). Computing with

these ideas, we see

$$\sum_{([u,v],x) \in \pi'} [F(v) - F(u)] = F(b) - F(a) - \sum_{i=1}^k [F(v_i) - F(u_i)]$$

and

$$\sum_{([u,v],w) \in \pi'} f(w)(v-u) = \sum_{([u,v],w) \in \pi''} f(w)(v-u) - \sum_{i=1}^k \left(\sum_{([u,v],w) \in \pi_i} f(w)(v-u) \right).$$

Putting these together with the estimates (6.5) and (6.6) we obtain

$$\begin{aligned} \left| \sum_{([u,v],x) \in \pi'} [[F(v) - F(u)] - f(x)(v-u)] \right| &\leq \left| F(b) - F(a) - \sum_{([u,v],x) \in \pi''} f(x)(v-u) \right| \\ &+ \sum_{i=1}^k \left| [F(v_i) - F(u_i)] - \sum_{([u,v],x) \in \pi_i} f(x)(v-u) \right| < \varepsilon/4 + k(\varepsilon/(4k)) = \varepsilon/2. \end{aligned}$$

Let us emphasize what we now see: if π' is any subset of π we have obtained this inequality:

$$\left| \sum_{([u,v],w) \in \pi'} [F(v) - F(u) - f(w)(v-u)] \right| < \varepsilon/2.$$

To complete the proof let

$$\pi^+ = \{([u,v],w) \in \pi : F(v) - F(u) - f(w)(v-u) \geq 0\}$$

and

$$\pi^- = \{([u,v],w) \in \pi : F(v) - F(u) - f(w)(v-u) < 0\}.$$

Then

$$\begin{aligned} &\sum_{([u,v],w) \in \pi^+} |F(v) - F(u) - f(w)(v-u)| \\ &= \sum_{([u,v],w) \in \pi^+} [F(v) - F(u) - f(w)(v-u)] < \varepsilon/2 \end{aligned}$$

and

$$\begin{aligned} & \sum_{([u,v],w) \in \pi^-} |F(v) - F(u) - f(w)(v-u)| \\ &= \sum_{([u,v],w) \in \pi^-} -[F(v) - F(u) - f(w)(v-u)] < \varepsilon/2. \end{aligned}$$

Adding the two inequalities proves (6.4). ■

Lemma 6.13 (HK integral is equivalent to the calculus integral) *Suppose that the function $f : [a, b] \rightarrow \mathbb{R}$ is Henstock-Kurzweil integrable on a compact interval $[a, b]$ and that F is its indefinite Henstock-Kurzweil integral. Then*

1. F is uniformly continuous on $[a, b]$.
2. F is absolutely continuous on (a, b) .
3. $F'(x) = f(x)$ for almost every x in (a, b) .

Proof. Let F be extended to allow $F(x) = F(a)$ for $x < a$ and $F(x) = F(b)$ for $x > b$ and let f be extended so that $f(x) = 0$ if x is not in the interval $[a, b]$. We can make use of the Henstock criterion proved in Lemma 715 and this extended version of F and f to claim that, for every $\varepsilon > 0$ there exists a full cover β (of the real line) such that

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u) - f(w)(v-u)| < \varepsilon,$$

for every subpartition π contained in β .

Write

$$h([u, v], w) = |F(v) - F(u) - f(w)(v-u)|$$

and observe that, in the language of Section 5.6.1, this function h has zero variation. Consequently it has a zero derivative almost everywhere. But at every point w at which h has a zero derivative, $F'(w) = f(w)$. In particular $F'(x) = f(x)$ for almost every x in (a, b) .

Take now any set N of measure zero. Write, for every integer $j = 1, 2, 3, \dots$

$$N_j = \{x \in N : j-1 \leq |f(x)| < j\}.$$

Each of these is a measure zero set and so there is a full cover β_j of N_j so that

$$\left| \sum_{([u,v],w) \in \pi} (v-u) \right| < \varepsilon j^{-1}$$

and, hence,

$$\sum_{([u,v],w) \in \pi} |f(w)|(v-u) < \varepsilon$$

whenever π is a subpartition chosen from β_j . The covering relation $\beta' = \beta \cap \beta_j$ is a full cover of N_j and

$$\begin{aligned} \sum_{([u,v],w) \in \pi} |F(v) - F(u)| &\leq \sum_{([u,v],w) \in \pi} |F(v) - F(u) - f(w)(v-u)| \\ &+ \sum_{([u,v],w) \in \pi} |f(w)|(v-u) < 2\varepsilon. \end{aligned}$$

It follows that F has zero variation on N_j . This is true for each $j = 1, 2, 3, \dots$ and so F has zero variation on N . This is true for any set of measure zero. Consequently F is absolutely continuous on $(-\infty, \infty)$. In particular it is continuous at each point and so also uniformly continuous on $[a, b]$. ■

6.3 Elementary properties of the integral

All of our elementary properties of the integral are anticipated by the naive calculus integral which shares all the same properties in somewhat weaker forms. Our interest here is that these same properties now hold under very general hypotheses. The reader should be able to construct proofs that use either the descriptive version of the calculus integral or the Henstock-Kurzweil version.

6.3.1 Integration and order

Theorem 6.14 *Suppose that f, g are both integrable on a compact interval $[a, b]$ and that $f(x) \leq g(x)$ for almost every x in that interval. Then*

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

6.3.2 Integration of linear combinations

Theorem 6.15 Suppose that f, g are both integrable on a compact interval $[a, b]$. Then so too is any linear combination $rf + sg$ and

$$\int_a^b [rf(x) + sg(x)] dx = r \left(\int_a^b f(x) dx \right) + s \left(\int_a^b g(x) dx \right).$$

6.3.3 Integrability on subintervals

Theorem 6.16 Suppose that f is integrable on a compact interval $[a, b]$. Then f is integrable on any compact subinterval of $[a, b]$.

6.3.4 Additivity

Theorem 6.17 If f is integrable on each of the intervals $[a, b]$, $[b, c]$, and $[a, c]$ then

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx.$$

6.3.5 Change of variable

Let $\phi : [a, b] \rightarrow \mathbb{R}$ be a strictly increasing differentiable function. We would expect from elementary formulas of the calculus that

$$\int_{\phi(a)}^{\phi(b)} f(x) dx = \int_a^b f(\phi(t))\phi'(t) dt.$$

If f is itself everywhere a derivative then this could be justified. If f is assumed only to be integrable then a different proof, using ϕ to map full covers and partitions, is needed.

Theorem 6.18 (Change of variable) Let $\phi : \mathbb{R} \rightarrow \mathbb{R}$ be a strictly increasing, differentiable function. If $f : \mathbb{R} \rightarrow \mathbb{R}$ is integrable on $[\phi(a), \phi(b)]$ then

$$\int_{\phi(a)}^{\phi(b)} f(x) dx = \int_a^b f(\phi(t))\phi'(t) dt.$$

Proof. Let $\varepsilon > 0$ and define β to be the collection of all pairs $([x, y], z)$ subject only to the conditions that

$$\left| \frac{\phi(y) - \phi(x)}{y - x} - \phi'(z) \right| < \frac{\varepsilon}{2(b-a)(1 + |f(\phi(z))|)}.$$

Since ϕ is everywhere differentiable this is a full cover. Note that we can write $\phi(y) - \phi(x)$ also as $\lambda(J)$ where $J = \phi([x, y])$ is just the compact interval that ϕ maps $[x, y]$ onto.

Write

$$\beta'_1 = \{(\phi([x, y]), \phi(x)) : ([x, y], z) \in \beta_1\}$$

and check that β'_1 is also a full cover. Observe that elements $(J, x) = (\phi([x, y]), \phi(x))$ of β'_1 must satisfy

$$|f(\phi(x))\lambda(\phi([x, y])) - f(\phi(x))\phi'(x)\lambda([x, y])| < \varepsilon\lambda([x, y])/2(b-a).$$

The expression $f(\phi(t))\lambda(\phi([x, y]))$ here is better viewed as $f(x)\lambda(J)$.

Choose a full cover β'_2 so that

$$\left| \int_{\phi(a)}^{\phi(b)} f(x) dx - \sum_{(J, x) \in \pi'} f(x)\lambda(J) \right| < \varepsilon/2$$

for all partitions $\pi' \subset \beta'_2$ of the interval $[\phi(a), \phi(b)]$. Write β_2 for the collection of all (I, x) for which $(I, x) = (\phi(J), \phi(t))$ for some $(J, t) \in \beta'_2$. This is a full cover of $[a, b]$.

Write $\beta = \beta_1 \cap \beta_2$. Check that β is a full cover of $[a, b]$ and check that

$$\left| \int_{\phi(a)}^{\phi(b)} f(x) dx - \sum_{(I, x) \in \pi} f(\phi(x))\phi'(x)\lambda(I) \right| < \varepsilon$$

for all partitions $\pi \subset \beta$ of the interval $[a, b]$. An appeal to the first Cauchy criterion then completes the proof. ■

6.3.6 Integration by parts

Integration by parts formula:

$$\int_a^b F(x)G'(x) dx = F(b)G(b) - F(a)G(b) - \int_a^b F'(x)G(x) dx. \quad (6.7)$$

The formula can be derived from the product rule for derivatives:

$$\frac{d}{dx}(F(x)G(x)) = F(x)G'(x) + F'(x)G(x)$$

which holds at any point where both functions are differentiable. One must then give strong enough hypotheses that the function $F(x)G(x)$ is an indefinite integral for the function

$$F(x)G'(x) + F'(x)G(x)$$

in the sense needed for our integral.

The most general statement is the following: if f and g are both integrable on $[a, b]$ and F and G are their indefinite integrals on that interval then $Fg + fG$ is integrable on $[a, b]$ and

$$\int_a^b (F(x)g(x) + f(x)G(x)) dx = F(b)G(b) - F(a)G(b).$$

In particular the usual formula (6.7) holds if and only if one of the two integrals in that statement exists. The proof is easiest to deduce from the Stieltjes version

$$\int_a^b F(x) dG(x) + G(x) dF(x) = F(b)G(b) - F(a)G(b) \quad (6.8)$$

that we will study in Chapter 8. The reader may wish to try, however, to prove it directly.

Remark: For the Lebesgue integral of Chapter 7 the integration by parts formula is available but not quite as straightforward. It is possible that $Fg + fG$ is integrable on $[a, b]$ but that only one of Fg and fG is Lebesgue integrable (i.e., absolutely integrable) on $[a, b]$. For example take $F(x) = x$ and $G(x) = x \cos x^{-2}$ on $[0, 1]$. It is also possible neither is Lebesgue integrable: take $F(x) = x^{1/2} \sin x^{-1}$ and $G(x) = x^{1/2} \cos x^{-1}$.

6.3.7 Derivative of the integral

If f is integrable on an interval $[a, b]$ then the formula

$$\frac{d}{dx} \int_a^x f(t) dt = f(x)$$

holds at almost every point in (a, b) . This is merely by definition. To make a claim, however, at some particular point the following simple observation is useful. We have seen it before in our study of the naive calculus integral. The proof

is the same.

Theorem 6.19 Let $f : [a, b] \rightarrow \mathbb{R}$ be an integrable function on the interval $[a, b]$.

Let

$$F(t) = \int_a^t f(x) dx \quad (a \leq t \leq b).$$

Assume that $x_0 \in [a, b]$ is a point of continuity of f . Then

1. If $a < x_0 < b$ then $F'(x_0) = f(x_0)$.
2. If $a = x_0$ then the right hand derivative $F'_+(a) = f(a)$.
3. If $x_0 = b$ then the left hand derivative $F'_-(b) = f(b)$.

Proof. Let x_0 be a point of continuity of f and let $\varepsilon > 0$. Then there is a $\delta > 0$ so that $|f(x) - f(x_0)| < \varepsilon$ if $|x - x_0| < \delta$ and $x \in [a, b]$. Let $[u, v] \subset [a, b]$ be any interval that contains x_0 and has length less than δ . Simply compute

$$\begin{aligned} \left| \int_u^v f(x) dx - f(x_0)(v - u) \right| &= \left| \int_u^v f(x) dx - \int_u^v f(x_0) dx \right| \\ &\leq \int_u^v |f(x) - f(x_0)| dx \leq \varepsilon(v - u). \end{aligned}$$

From this the conclusions of the theorem are easy to check. ■

6.3.8 Null functions

A function is a null function if it is equal to zero at every point with only a small set of exceptions. It is immediately clear that every null function has a constant indefinite integral. Thus the following statements are obvious.

Theorem 6.20 Let $f : [a, b] \rightarrow \mathbb{R}$ be a null function. Then f is integrable on $[a, b]$

and

$$\int_a^b f(x) dx = 0.$$

6.3. ELEMENTARY PROPERTIES OF THE INTEGRAL

Theorem 6.21 Let $f : [a, b] \rightarrow \mathbb{R}$ be an integrable function with the property that

$$\int_c^d f(x) dx = 0 \text{ for all } [c, d] \subset [a, b].$$

Then f is a null function.

Corollary 6.22 Let $f : [a, b] \rightarrow \mathbb{R}$ be a nonnegative integrable function with the property that

$$\int_a^b f(x) dx = 0.$$

Then f is a null function.

6.3.9 Monotone convergence theorem

The formula

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \left(\lim_{n \rightarrow \infty} f_n(x) \right) dx$$

is extremely useful but not generally valid. If the sequence of integrable functions $\{f_n\}$ is monotone then this does hold.

Theorem 6.23 (Monotone convergence theorem) Let $f_n : [a, b] \rightarrow \mathbb{R}$ ($n = 1, 2, 3, \dots$) be a nondecreasing sequence of integrable functions and suppose that

$$f(x) = \lim_{n \rightarrow \infty} f_n(x)$$

for almost every x in $[a, b]$. Then

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx. \quad (6.9)$$

In particular, if the limit exists and is finite the function f is integrable on $[a, b]$ and the identity (6.9) holds. If the limit is infinite then the function f is not integrable but the integral is determined and

$$\int_a^b f(x) dx = \infty.$$

Here we are using the ideas from Section 6.2.2 that allow us to express an integral as infinite. This was not available to us in our study of the calculus integral but the Henstock-Kurzweil theory of upper and lower integrals allowed this. The proof of Theorem 6.23 is given in Section 6.3.11 below.

6.3.10 Summing inside the integral

We establish here that the summation formula

$$\int_a^b \left(\sum_{n=1}^{\infty} f_n(x) \right) dx = \sum_{n=1}^{\infty} \left(\int_a^b f_n(x) dx \right)$$

is possible for nonnegative integrable functions.

Theorem 6.24 (summing inside the integral) *Let $f_n : [a, b] \rightarrow \mathbb{R}$ ($n = 1, 2, 3, \dots$) be a sequence of nonnegative integrable functions and suppose that*

$$f(x) = \sum_{n=1}^{\infty} f_n(x)$$

for almost every x . Then

$$\int_a^b f(x) dx = \sum_{n=1}^{\infty} \left(\int_a^b f_n(x) dx \right).$$

In particular, if the series converges the function f is integrable on $[a, b]$ and the identity (6.12) holds. If the series diverges then the function f is not integrable but the integral is determined and

$$\int_a^b f(x) dx = \infty.$$

The proof is obtained from the two lemmas given in Section 6.3.11 below.

6.3.11 Two convergence lemmas

The monotone convergence theorem and the formula for summing inside the integral are directly related by the following observation. If

$$f_1(x) \leq f_2(x) \leq f_3(x) \leq \dots$$

and

$$\lim_{n \rightarrow \infty} f_n(x) = f(x)$$

then

$$f(x) = f_1(x) + \sum_{n=1}^{\infty} (f_n(x) - f_{n-1}(x))$$

expresses f as the sum of a series. Thus it is enough to prove Theorem 6.24. This is obtained from the following two lemmas.

Lemma 6.25 Suppose that $f, f_1, f_2, \dots$ is a sequence of nonnegative functions defined on a compact interval $[a, b]$. If, for almost every x

$$f(x) \geq \sum_{n=1}^{\infty} f_n(x),$$

then

$$\int_a^b f(x) dx \geq \sum_{n=1}^{\infty} \left(\int_a^b f_n(x) dx \right). \quad (6.10)$$

Proof. We can assume that the inequality assumed is valid for every x ; simply redefine $f_n(x) = 0$ for those points in the null set where the inequality doesn't work. The resulting functions will have the same lower integrals as f_n .

Let $\varepsilon > 0$. Take any integer N and choose full covers β_n ($n = 1, 2, \dots, N$) so that all the Riemann sums¹

$$\sum_{\pi} f_n(w)(v-u) \geq \int_a^b f_n(x) dx - \varepsilon 2^{-n}$$

whenever $\pi \subset \beta_n$ is a partition of $[a, b]$. (If the integrals here are not finite then there is nothing to prove, since both sides of the inequality (6.10) will be infinite.)

Let

$$\beta = \bigcap_{n=1}^N \beta_n.$$

This too is a full cover, one that is contained in all of the others.

Take any partition of $[a, b]$ with $\pi \subset \beta$, and compute

$$\begin{aligned} \sum_{\pi} f(w)(v-u) &\geq \sum_{\pi} \left(\sum_{n=1}^N f_n(w)(v-u) \right) = \sum_{n=1}^N \left(\sum_{\pi} f_n(w)(v-u) \right) \geq \\ &\sum_{n=1}^N \left(\int_a^b f_n(x) dx - \varepsilon 2^{-n} \right). \end{aligned}$$

This gives a lower bound for all Cauchy sums and hence, since ε is arbitrary, shows that

$$\int_a^b f(x) dx \geq \sum_{n=1}^N \left(\int_a^b f_n(x) dx \right).$$

As this is true for all N the inequality (6.10) must follow. ■

¹We simplify our notation for Riemann sums a bit by replacing

$$\sum_{([u,v],w) \in \pi} f(w)(v-u) \text{ by } \sum_{\pi} f(w)(v-u).$$

Lemma 6.26 Suppose that $f, f_1, f_2, \dots$ is a sequence of nonnegative functions defined on a compact interval $[a, b]$. If, for almost every x

$$f(x) \leq \sum_{n=1}^{\infty} f_n(x),$$

then

$$\int_a^b f(x) dx \leq \sum_{n=1}^{\infty} \left(\int_a^b f_n(x) dx \right). \quad (6.11)$$

Proof. As before, we can assume that the inequality assumed is valid for every x ; simply redefine $f(x) = 0$ for those points in the null set where the inequality doesn't work. The resulting function will have the same integral and same upper integral as f .

This lemma is similar to the preceding one, but requires a bit of bookkeeping and a new technique with the covers. Let $t < 1$ and choose for each $x \in [a, b]$ the first integer $N(x)$ so that

$$t f(x) \leq \sum_{n=1}^{N(x)} f_n(x).$$

Choose, again and using the same ideas as before, full covers β_n ($n = 1, 2, \dots$) so that $\beta_1 \supset \beta_2 \supset \beta_3 \dots$ and all Riemann sums²

$$\sum_{\pi} f_n(w)(v-u) \leq \int_a^b f_n(x) dx + \epsilon 2^{-n}$$

whenever $\pi \subset \beta_n$ is a partition of $[a, b]$. (Again, if the integrals here are not finite then there is nothing to prove, since the larger side of the inequality (6.11) will be infinite.)

Let

$$E_n = \{x \in [a, b] : N(x) = n\}.$$

²As before, we simplify our notation for Riemann sums by replacing

$$\sum_{([u,v],w) \in \pi} f(w)(v-u) \text{ by } \sum_{\pi} f(w)(v-u).$$

We use these sets to carve up the covering relations. Write

$$\beta_n[E_n] = \{([u, v], w) \in \beta_n : w \in E_n\}.$$

There must be a full cover β so that

$$\beta[E_n] \subset \beta_n[E_n]$$

for all $n = 1, 2, 3, \dots$.

Take any partition of $[a, b]$ with $\pi \subset \beta$. Let N be the largest value of $N(x)$ for the finite collection of pairs $(I, x) \in \pi$. We need to carve the partition π into a finite number of disjoint subsets by writing, for $j = 1, 2, 3, \dots, N$,

$$\pi_j = \{([u, v], w) \in \pi : w \in E_j\}$$

and

$$\sigma_j = \pi_j \cup \pi_{j+1} \cup \dots \cup \pi_N.$$

for integers $j = 1, 2, 3, \dots, N$. Note that

$$\sigma_j \subset \beta_j$$

and that

$$\pi = \pi_1 \cup \pi_2 \cup \dots \cup \pi_N.$$

Check the following computations, making sure to use the fact that for $x \in E_i$,

$$t f(x) \leq f_1(x) + f_2(x) + \dots + f_i(x).$$

$$\begin{aligned} \sum_{\pi} t f(w)(v-u) &= \sum_{i=1}^N \sum_{\pi_i} t f(w)(v-u) \\ &\leq \sum_{i=1}^N \sum_{\pi_i} (f_1(w) + f_2(w) + \dots + f_i(w))(v-u) \\ &= \sum_{j=1}^N \left(\sum_{\sigma_j} f_j(w)(v-u) \right) \leq \\ &\sum_{j=1}^N \left(\int_a^b f_j(x) dx + \varepsilon 2^{-j} \right) \leq \sum_{j=1}^{\infty} \left(\int_a^b f_j(x) dx \right) + \varepsilon. \end{aligned}$$

This gives an upper bound for all Cauchy sums and hence, since ε is arbitrary, shows that

$$\int_a^b t f(x) dx \leq \sum_{n=1}^{\infty} \left(\int_a^b f_n(x) dx \right).$$

As this is true for all $t < 1$ the inequality (6.11) must follow too. ■

Exercises

Exercise 658 Give an example to show that it is possible that $\int_a^b f(x) dx = \infty$ in Theorem 6.24. □

Exercise 659 Give an example to show that it is possible for the Theorem 6.24 to fail if we drop the assumption that the functions are nonnegative in the theorem. □

Exercise 660 Let $f_n : [a, b] \rightarrow \mathbb{R}$ ($n = 1, 2, 3, \dots$) be a sequence of absolutely integrable functions and suppose that

$$\sum_{n=1}^{\infty} |f_n(x)| < \infty$$

for almost every x and that

$$\sum_{n=1}^{\infty} \left(\int_a^b |f_n(x)| dx \right) < \infty.$$

Then show that

$$f(x) = \sum_{n=1}^{\infty} f_n(x)$$

is finite for almost every x in $[a, b]$, is absolutely integrable, and that

$$\int_a^b f(x) dx = \sum_{n=1}^{\infty} \left(\int_a^b f_n(x) dx \right).$$

□

6.4 Equi-integrability

We can use Definition 6.6 to describe a uniform version of integrability that is useful in discussions of the convergence of sequences of integrable functions.

Definition 6.27 (equi-integrability) Suppose that $\{f_n\}$ is a sequence of integrable functions defined at every point of a compact interval $[a, b]$. Then $\{f_n\}$ is said to be equi-integrable on $[a, b]$ if, for every $\varepsilon > 0$, there is a full cover β of $[a, b]$ so that

$$\left| \int_a^b f_n(x) dx - \sum_{[u,v],w \in \pi} f_n(w)(v-u) \right| < \varepsilon$$

whenever π is a partition of the interval $[a, b]$ chosen from β .

Uniform convergence is a sufficient condition for equi-integrability, but the condition itself is much more general.

Lemma 6.28 Suppose that $\{f_n\}$ is a sequence of integrable functions defined at every point of a compact interval $[a, b]$ and that $\{f_n\}$ is uniformly convergent on $[a, b]$. Then $\{f_n\}$ is equi-integrable on $[a, b]$.

Equi-integrability along with pointwise convergence gives a simply stated criterion for taking the limit inside the integral.

Theorem 6.29 Suppose that $\{f_n\}$ is a sequence of equi-integrable functions defined at every point of a compact interval $[a, b]$ and that $\{f_n\}$ is pointwise convergent on $[a, b]$ to a function f . Then f is integrable on $[a, b]$ and

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx. \quad (6.12)$$

Chapter 7

Lebesgue's Integral

Lebesgue's program is the construction of the value of the integral

$$\int_a^b f(x) dx$$

directly from the measure and the values of the function f in the integral. Our formal definition of the integral *appears* to do this. Since full covers are not themselves, in general, constructible from the function being integrated we cannot claim that our integral is constructed in the sense Lebesgue intends.

For his program he invented the integral as a heuristic device, imagined what properties it should possess and then went about discovering how to construct it based on this fiction. At the end he then had to take his construction as the definition itself. For us to follow the same program is much easier: we have an integral, we know many of its properties, and we can use this information to construct it.

This chapter presents an introduction to Lebesgue's methods, but backwards in a sense from conventional presentations. We already have a formal definition of the integral, so we do not need to define an integral by Lebesgue's method. We need to show how to construct the value of an object $\int_a^b f(x) dx$ that we have already defined by other means.

7.1 The Lebesgue integral

The Lebesgue integral is a special case of the general calculus integral. It is not merely a special case, but certainly the most important special case.

Definition 7.1 *Let f be a function defined almost everywhere on an interval $[a, b]$. Then f is said to be Lebesgue integrable if f is absolutely integrable, i.e., if both f and $|f|$ are integrable on $[a, b]$.*

Functions that are integrable but not Lebesgue integrable are said to be nonabsolutely integrable. The theory of such functions is less powerful and more delicate than the theory of the Lebesgue integrable functions. There are also fewer applications. We return to this topic in Chapter 9.

7.2 Lebesgue measure

We define the following three versions of Lebesgue measure (similar to the three versions of a measure zero set) for a set $E \subset \mathbb{R}$:

- $\lambda(E) = \inf\{\lambda(G) : G \text{ open and } G \supset E\}.$

- $\lambda^*(E) = \inf_{\beta} \left(\sup_{\pi \subset \beta} \sum_{([u,v],w) \in \pi} (v - u) \right)$

where the infimum is taken over all full covers β of the set E and $\pi \subset \beta$ is an arbitrary subpartition.

- $\lambda_*(E) = \inf_{\beta} \left(\sup_{\pi \subset \beta} \sum_{([u,v],w) \in \pi} (v - u) \right)$

where the infimum is taken over all fine covers β of the set E and $\pi \subset \beta$ is an arbitrary subpartition.

The first of these λ is Lebesgue's original version of his measure. We have already (in Section 5.2.1) defined the Lebesgue measure of open sets. This definition extends that, by a simple infimum, to all sets. The definition of the full measure λ^* is closely related to the integral.

Lemma 7.2 *Let E be a set of real numbers contained in an interval $[a, b]$. Then*

$$\lambda^*(E) = \int_a^b \chi_E(x) dx.$$

The three definitions are equivalent, a fact which is proved as the Vitali covering theorem in Section 7.3 below.

7.2.1 Basic property of Lebesgue measure

Theorem 7.3 *Lebesgue measure λ is a nonnegative real-valued set function defined for all sets of real numbers that is a measure^a on $\mathbb{R}$, i.e., it has the following properties:*

1. $\lambda(\emptyset) = 0$.
2. For any sequence of sets $E, E_1, E_2, E_3, \dots$ for which

$$E \subset \bigcup_{n=1}^{\infty} E_n$$

the inequality

$$\lambda(E) \leq \sum_{n=1}^{\infty} \lambda(E_n)$$

must hold.

^aMost authors would call this an *outer measure*.

This result is often described by the following language that splits the property (2) in two parts:

Subadditivity: $\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} \lambda(E_n).$

Monotonicity: $\lambda(A) \leq \lambda(B)$ if $A \subset B$.

Since we have three representations of the Lebesgue measure, as λ , λ^* , or as λ_* we can prove this using any one of the three. The exercises ask for all three; any one would suffice in view of the Vitali covering theorem proved in the next section.

Exercises

Exercise 661 Establish the identity

$$\lambda^*(E) = \overline{\int_a^b \chi_E(x) dx}$$

for any set $E \subset [a, b]$.

Answer ☐

Exercise 662 Prove that λ is a measure in the sense of Theorem 7.3.

Answer ☐

Exercise 663 Prove that λ^* is a measure in the sense of Theorem 7.3.

☐

Exercise 664 Prove that λ_* is a measure in the sense of Theorem 7.3.

☐

Exercise 665 Prove that $\lambda(A) < t$ if and only if there is an open set G that contains all but countably many points of A and for which $\lambda(G) < t$.

☐

7.3 Vitali covering theorem

These three measures are identical and we can use any version. The identity $\lambda = \lambda^* = \lambda_*$ is Vitali's theorem, although his theorem is normally expressed in different language.¹

The proof is just a bit more difficult than the proof of the narrower version, the mini-Vitali theorem given in Section 5.5, where we showed that sets of measure zero were equivalent to both full null and fine null sets.

Theorem 7.4 (Vitali Covering Theorem) $\lambda = \lambda^* = \lambda_*$.

¹The language here, will no doubt, shock some traditionalists for whom it appears to suggest Lebesgue inner and outer measure. But this has nothing to do with inner/outer measure. The measures λ^* and λ_* are those derived from full and fine covers.

7.3.1 Classical version of Vitali's theorem

Vitali's covering theorem asserts that the measure of an arbitrary set can be determined from full and fine covers of that set. The basic computation about fine covers is the following lemma, known as the classical version of Vitali's theorem.

Lemma 7.5 (Vitali covering theorem) *Let β be a fine cover of a bounded set E and suppose that $\varepsilon > 0$. Then there must exist a subpartition $\pi \subset \beta$ for which*

$$\lambda \left(E \setminus \bigcup_{([u,v],w) \in \pi} [u,v] \right) < \varepsilon. \quad (7.1)$$

Proof. For the proof of this theorem we need only one simple fact (Exercise 665) about the Lebesgue measure $\lambda(E)$ of a real set A :

$\star \lambda(A) < \varepsilon$ if and only if there is an open set G containing all but countably many points of A and for which $\lambda(G) < \varepsilon$.

Thus the proof is really about open sets. Indeed in our proof we use only the Lebesgue measure of open sets and several covering lemmas.

The proof is just a repeated application of Lemma 5.17. Since E is bounded there is an open set U_1 containing E for which $\lambda(U_1) < \infty$. If $\lambda(U_1) < \varepsilon$ then, since $E \subset U_1$, $\lambda(E) < \varepsilon$ and there is nothing more to prove: take $\pi = \emptyset$ and the statement (7.1) is satisfied. If $\lambda(U_1) \geq \varepsilon$ we start our process.

We prune β by the open set U_1 : define $\beta_1 = \beta(U_1)$. Note that this, too, is a fine cover of E . Set

$$G_1 = \bigcup_{([u,v],w) \in \beta_1} (u,v).$$

Then G_1 is an open set and $g_1 = \lambda(G_1) < \lambda(U_1)$ is finite. We know from Lemma 5.16, that G_1 covers all of E except for a countable set. [We shall ignore countable sets in this proof, to keep the bookkeeping simple]. By Lemma 5.17 there must exist a subpartition $\pi_1 \subset \beta_1$ for which

$$U_2 = G_1 \setminus \bigcup_{([u,v],w) \in \pi_1} [u,v]$$

is an open subset of G_1 and

$$\lambda(U_2) \leq 5g_1/6 \leq 5\lambda(U_1)/6.$$

Define

$$E_1 = E \setminus \bigcup_{([u,v],w) \in \pi_1} [u, v].$$

If $\lambda(U_2) < \varepsilon$ then $\lambda(E_1) < \varepsilon$. This is because U_2 is an open set containing all of E_1 except possibly some countable set; thus $\star$ stated above implies that $\lambda(E_1) < \varepsilon$. But if $\lambda(E_1) < \varepsilon$ the process can stop: take $\pi = \pi_1$ and the statement (7.1) is satisfied.

If $\lambda(U_2) \geq \varepsilon$ we continue our process. Define $\beta_2 = \beta(U_2)$ and note that this is a fine cover of E_1 (i.e., the points in E not already handled by the subpartition π_1 or the countably many points of E discarded in the first stage of our proof).

Set

$$G_2 = \bigcup_{([u,v],w) \in \beta_2} (u, v).$$

Then G_2 is an open set and

$$g_2 = \lambda(G_2) \leq \lambda(U_2).$$

As before, we know from Lemma 5.16, that G_2 covers all of E_1 except for a countable set. [We are ignoring countable sets in this proof, throw these points away].

Again applying Lemma 5.17, we find a subpartition $\pi_2 \subset \beta_2$ for which

$$U_3 = G_2 \setminus \bigcup_{([u,v],w) \in \pi_2} [u, v]$$

is an open subset of G_2 and $\lambda(U_3) \leq 5g_2/6$. Define

$$\begin{aligned} E_2 &= E_1 \setminus \bigcup_{([u,v],w) \in \pi_2} [u, v] \\ &= E \setminus \bigcup_{([u,v],w) \in \pi_1 \cup \pi_2} [u, v]. \end{aligned}$$

If $\lambda(U_3) < \varepsilon$ then $\lambda(E_2) < \varepsilon$. This because U_3 is an open set containing all of E_2 except possibly some countable set; thus $\star$ stated above implies that $\lambda(E_2) < \varepsilon$. But if $\lambda(E_2) < \varepsilon$ the process can stop: take $\pi = \pi_1 \cup \pi_2$ and the statement

(7.1) is satisfied. [Be sure to check that the intervals from π_1 have been arranged to be disjoint from the intervals in π_2 .]

This process is continued, inductively, until it stops. It certainly *must* stop since

$$\lambda(U_{k+1}) < \frac{5}{6}\lambda(U_k) \leq \cdots \leq \left(\frac{5}{6}\right)^k \lambda(U_1)$$

so that eventually $\lambda(U_{k+1}) < \varepsilon$ and $\lambda(E_k) < \varepsilon$. Take

$$\pi = \pi_1 \cup \pi_2 \cup \dots \pi_k$$

and the statement (7.1) is satisfied. ■

7.3.2 Proof that $\lambda = \lambda^* = \lambda_*$.

The inequality

$$\lambda_* \leq \lambda^* \leq \lambda$$

is trivial. First of all, any full cover is also a fine cover so that $\lambda_* \leq \lambda^*$ must be true. Second, if $\lambda(E) < t$ there is an open set G containing E for which it is also true that $\lambda(G) < t$. But then we can define a covering relation β to consist of all pairs $([u, v], w)$ provided $w \in [u, v] \subset G$. This is a full cover of E . Note that

$$\sum_{([u, v], w) \in \pi} (v - u) \leq \lambda(G) < t$$

whenever $\pi \subset \beta$ is an arbitrary subpartition. It follows that $\lambda^*(E) < t$. As this is true for all t ,

$$\lambda^*(E) \leq \lambda(E).$$

Finally, then, Lemma 7.5 completes the proof. Let β be any fine cover of a bounded set E and suppose that $\varepsilon > 0$. Then there must exist a subpartition $\pi \subset \beta$ for which

$$\lambda \left(E \setminus \bigcup_{([u, v], w) \in \pi} [u, v] \right) < \varepsilon. \quad (7.2)$$

In particular, using subadditivity measure property of λ ,

$$\begin{aligned}\lambda(E) &\leq \lambda\left(E \setminus \bigcup_{([u,v],w) \in \pi} [u,v]\right) + \sum_{([u,v],w) \in \pi} \lambda([u,v]) \\ &< \sum_{([u,v],w) \in \pi} (v-u) + \varepsilon.\end{aligned}$$

So, since this is true for any fine cover of E ,

$$\lambda(E) \leq \lambda_*(E) + \varepsilon.$$

It follows that $\lambda(E) \leq \lambda_*(E)$ for all bounded sets E .

That establishes the identity $\lambda = \lambda^* = \lambda_*$ for all bounded sets. The extension to unbounded sets can be accomplished with the standard measure properties.

7.4 Density theorem

As an application of the Vitali covering theorem we prove the density theorem. This asserts that for an arbitrary set E almost every point is a point of density, a point x where

$$\frac{\lambda(E \cap [u, v])}{\lambda([u, v])} \rightarrow 1$$

as $[u, v]$ shrinks to x .

Theorem 7.6 *Almost every point of an arbitrary set E is a point of density.*

Proof. To define this with a bit more precision write

$$\underline{d}(E, x) = \sup_{\delta > 0} \inf \left\{ \frac{\lambda(E \cap [u, v])}{\lambda([u, v])} : u \leq x \leq v, 0 < v - u < \delta \right\}.$$

This is called the *lower density* of E at x . The theorem asserts that

$$\underline{d}(E, x) = 1$$

at almost every point x of E .

We may assume that E is bounded. Take any $\alpha < 1$ and define

$$E_\alpha = \{x \in E : \underline{d}(E, x) < \alpha\}$$

and

$$E' = \{x \in E : \underline{d}(E, x) < 1\}.$$

We show that E_α is necessarily a set of measure zero. It follows that E' is then a set of measure zero since evidently

$$E' = \bigcup_{n=1}^{\infty} E_{\frac{n}{n+1}}.$$

Fix $\alpha < 1$ and any open set G containing E_α , and define

$$\beta = \{([u, v], w) : u \leq x \leq v, \lambda(E \cap [u, v]) < \alpha \lambda([u, v])\}.$$

This is a fine cover of E_α , and since G is an open set containing E_α , the pruned relation $\beta(G)$ is also a fine cover of E_α . Let $\varepsilon > 0$. By the Vitali covering theorem (Lemma 7.5) there must exist a subpartition $\pi \subset \beta(G)$ for which

$$\lambda \left(E_\alpha \setminus \bigcup_{([u, v], w) \in \pi} [u, v] \right) < \varepsilon. \quad (7.3)$$

Now we simply compute, using subadditivity, that

$$\begin{aligned} \lambda(E_\alpha) &\leq \lambda \left(E_\alpha \setminus \bigcup_{([u, v], w) \in \pi} [u, v] \right) + \sum_{([u, v], w) \in \pi} \lambda(E_\alpha \cap [u, v]) \\ &\leq \varepsilon + \sum_{([u, v], w) \in \pi} \lambda(E \cap [u, v]) \\ &\leq \varepsilon + \alpha \sum_{([u, v], w) \in \pi} \lambda([u, v]) \leq \varepsilon + \alpha \lambda(G). \end{aligned}$$

We deduce that $\lambda(E_\alpha) \leq \lambda(G)$ for all such open sets G and hence that $\lambda(E_\alpha) \leq \alpha \lambda(E_\alpha)$. This is possible only if $\lambda(E_\alpha) = 0$. ■

7.5 Additivity

Lebesgue measure is *subadditive* in general on the union of two sets E_1 and E_2 . The subadditivity formula is

$$\lambda(A \cap (E_1 \cup E_2)) \leq \lambda(A \cap E_1) + \lambda(A \cap E_2)$$

We know that this same subadditivity formula holds for a sequence of sets $\{E_i\}$:

$$\lambda \left(A \cap \left(\bigcup_{i=1}^{\infty} E_i \right) \right) \leq \sum_{i=1}^{\infty} \lambda(A \cap E_i).$$

We now ask for conditions under which we can claim equality (not inequality). The additivity formula we wish to investigate is

$$\lambda \left(A \cap \left(\bigcup_{i=1}^{\infty} E_i \right) \right) = \sum_{i=1}^{\infty} \lambda(A \cap E_i)?$$

Our first observation is that this is possible if the sets $\{E_i\}$ are separated by open sets. This means merely that there exist open sets G_i and G_j that have no point in common, with $E_i \subset G_i$ and $E_j \subset G_j$. This is stronger than the requirement that E_i and E_j have no point in common. But note that two disjoint closed sets can always be separated in this fashion.

Lemma 7.7 *Let E_1 and E_2 be sets that are separated by open sets. Then, for any set A*

$$\lambda(A \cap (E_1 \cup E_2)) = \lambda(A \cap E_1) + \lambda(A \cap E_2).$$

Proof. Let us use the full version λ^* . We know that

$$\lambda^*(A \cap (E_1 \cup E_2)) \leq \lambda^*(A \cap E_1) + \lambda^*(A \cap E_2).$$

Let us prove the opposite direction. Let β be any full cover of $A \cap (E_1 \cup E_2)$. Select G_1 and G_2 , disjoint open sets containing E_1 and E_2 (respectively). Then $\beta(G_1 \cup G_2)$ is necessarily a full cover of $A \cap (E_1 \cup E_2)$. Note that $\beta(G_1)$ is a full cover of $A \cap E_1$ and that $\beta(G_2)$ is a full cover of $A \cap E_2$. If $t_1 < \lambda^*(A \cap E_1)$ and $t_2 < \lambda^*(A \cap E_2)$ then there must be subpartitions $\pi_1 \subset \beta(G_1)$ and $\pi_2 \subset \beta(G_2)$ with

$$\sum_{([u,v],w) \in \pi_1} (v - u) > t_1$$

7.5. ADDITIVITY

and

$$\sum_{([u,v],w) \in \pi_2} (v-u) > t_2.$$

It follows that β contains a subpartition $\pi = \pi_1 \cup \pi_2$ for which

$$\sum_{([u,v],w) \in \pi} (v-u) > t_1 + t_2.$$

From this we deduce that $\lambda^*(A \cap (E_1 \cup E_2)) > t_1 + t_2$. Then

$$\lambda^*(A \cap (E_1 \cup E_2)) \geq \lambda^*(A \cap E_1) + \lambda^*(A \cap E_2)$$

follows. ■

Corollary 7.8 *Let $E_1, E_2, E_3, \dots$ be a sequence of pairwise disjoint subsets of $\mathbb{R}$ and write*

$$E = \bigcup_{i=1}^{\infty} E_i.$$

Suppose that each pair of sets in the sequence are separated by open sets. Then, for any set A ,

$$\lambda(A \cap E) = \sum_{i=1}^{\infty} \lambda(A \cap E_i).$$

Proof. We know from the usual measure properties that

$$\lambda(A \cap E) \leq \sum_{i=1}^{\infty} \lambda(A \cap E_i).$$

We also know that

$$\lambda(A \cap (E_1 \cup E_2)) = \lambda(A \cap E_1) + \lambda(A \cap E_2).$$

An inductive argument would show, too, that for any $n > 1$,

$$\lambda(A \cap (E_1 \cup E_2 \cdots \cup E_n)) = \lambda(A \cap E_1) + \lambda(A \cap E_2) + \cdots + \lambda(A \cap E_n).$$

Thus, from the monotonicity property of measures,

$$\sum_{i=1}^n \lambda(A \cap E_i) \leq \lambda(A \cap E) \leq \sum_{i=1}^{\infty} \lambda(A \cap E_i).$$

From this the corollary evidently follows. ■

Corollary 7.9 *Let $E_1, E_2, E_3, \dots$ be a sequence of pairwise disjoint closed subsets of $\mathbb{R}$. Then, for any set A ,*

$$\lambda(A \cap E) = \sum_{i=1}^{\infty} \lambda(A \cap E_i).$$

To push the countable additivity one step further we use the previous corollary in a natural way. This looks like a highly technical lemma, but it is the basis and motivation for our definition of measurable sets and the theory is more natural than it might appear. The proof is left as an exercise; working through a proof should make it clear how and why the measurability definition in the next section works.

Lemma 7.10 *Let $E_1, E_2, E_3, \dots$ be a sequence of pairwise disjoint subsets of $\mathbb{R}$ and write*

$$E = \bigcup_{i=1}^{\infty} E_i.$$

Suppose that for every $\varepsilon > 0$ and for every n there is an open set G_n so that $E_n \setminus G_n$ is closed and so that $\lambda(G_n) < \varepsilon$. Then, for any set A ,

$$\lambda(A \cap E) = \sum_{i=1}^{\infty} \lambda(A \cap E_i).$$

7.6. MEASURABLE SETS

7.6 Measurable sets

7.6.1 Definition of measurable sets

Definition 7.11 An arbitrary subset E of $\mathbb{R}$ is measurable^a if for every $\varepsilon > 0$ there is an open set G with $\lambda(G) < \varepsilon$ and so that $E \setminus G$ is closed.

^aMost advanced courses will start with a different definition of measurable and later on show that this property used here is equivalent in certain settings. See Section 7.8.2 for the connections.

Thus a set is measurable if it is “almost closed.” Immediately from this definition we see that all closed sets are measurable and that all null sets are measurable. The definition is exactly designed to produce the following Theorem.

Theorem 7.12 Let $E_1, E_2, E_3, \dots$ be a sequence of pairwise disjoint measurable subsets of $\mathbb{R}$ and write

$$E = \bigcup_{i=1}^{\infty} E_i.$$

Then, for any set A ,

$$\lambda(A \cap E) = \sum_{i=1}^{\infty} \lambda(A \cap E_i).$$

Proof. This follows immediately from Lemma 7.12. ■

7.6.2 Properties of measurable sets

Theorem 7.13 The class of all measurable subsets of $\mathbb{R}$ forms a Borel family^a that contains all closed sets and all null sets.

^aThe definition of a Borel family is outlined in the proof.

Proof. The class of all measurable subsets of $\mathbb{R}$ forms a Borel family: it a collection of sets that is closed under the formation of unions and intersections of sequences of its members, and contains the complement of each of its members.

Here are the details of the proof. Items (3), (4), and (5) are specifically the requirements that the class of measurable sets forms a Borel family.

We prove that the family of all measurable sets has the following properties:

1. Every null set is measurable.
2. Every closed set is measurable.
3. If $E_1, E_2, E_3, \dots$ is a sequence of measurable sets then the union $\bigcup_{n=1}^{\infty} E_n$ is also measurable.
4. If $E_1, E_2, E_3, \dots$ is a sequence of measurable sets then the intersection $\bigcap_{n=1}^{\infty} E_n$ is also measurable.
5. If E is measurable then the complement $\mathbb{R} \setminus E$ is also measurable.

Items (1) and (2) are easy. Let us prove (5) first. Let E be measurable and E' is its complement. Let $\varepsilon > 0$ and choose an open set G_1 so that $E \setminus G_1$ is closed and $\lambda(G_1) < \varepsilon/2$. Let O be the complement of $E \setminus G_1$; evidently O is open.

First find an open set G_2 with $\lambda(G_2) < \varepsilon/2$ so that $O \setminus G_2$ is closed. [Simply display the component intervals of O , handle the infinite components first, and then a finite number of the bounded components.] Now observe that

$$E' \setminus (G_1 \cup G_2) = O \setminus G_2$$

is a closed set while $G_1 \cup G_2$ is an open set with measure smaller than ε . This verifies that E' is measurable.

Now check (e): let $\varepsilon > 0$ and choose open sets G_n so that $\lambda(G_n) < \varepsilon 2^{-n}$ and each $E_n \setminus G_n$ is closed. Observe that the set $G = \bigcup_{n=1}^{\infty} G_n$ is an open set for which

$$\lambda(G) \leq \sum_{n=1}^{\infty} \lambda(G_n) \leq \sum_{n=1}^{\infty} \varepsilon 2^{-n} = \varepsilon.$$

Finally

$$E' = E \setminus G = \bigcap_{n=1}^{\infty} (E_n \setminus G_n)$$

is closed.

For (4), write E'_n for the complementary set to E_n . Then the complement of the set $A = \bigcup_{n=1}^{\infty} E_n$ is the set $B = \bigcap_{n=1}^{\infty} E'_n$. Each E'_n is measurable by (5) and hence B is measurable by (d). The complement of B , namely the set A , is measurable by (5) again. ■

7.6.3 Increasing sequences of sets

If

$$E_1 \subset E_2 \subset E_3 \subset \dots$$

is an increasing sequence of sets then we would expect that

$$\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n \rightarrow \infty} \lambda(E_n).$$

This is particularly easy to prove if the sets are measurable. We show that this identity holds in general.

Theorem 7.14 Suppose that $\{E_n\}$ is an increasing sequence of sets. Then

$$\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n \rightarrow \infty} \lambda(E_n).$$

Proof. Suppose first that the sets are measurable. Then simply write $A_0 = \emptyset$ and $A_n = E_n \setminus E_{n-1}$ for each $n = 1, 2, 3, \dots$. Then these sets are also measurable and Lemma 7.12 shows us that

$$\begin{aligned} \lambda\left(\bigcup_{n=1}^{\infty} E_n\right) &= \lambda\left(\bigcup_{n=1}^{\infty} A_n\right) \\ &= \sum_{n=1}^{\infty} \lambda(A_n) = \sum_{n=1}^{\infty} (\lambda(E_n) - \lambda(E_{n-1})) = \lim_{n \rightarrow \infty} \lambda(E_n). \end{aligned}$$

Now we drop the assumption that the sets $\{E_n\}$ are measurable. Observe first that

$$\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) \geq \lim_{m \rightarrow \infty} \lambda(E_m)$$

merely because each set E_m is contained in this union.

To prove the opposite inequality, begin by choosing measurable sets $H_n \supset E_n$ with the same measures, i.e., so that $\lambda(E_n) = \lambda(H_n)$. (For example, start with a sequence of open sets G_{nm} containing E_n with $\lambda(E_n) \leq \lambda(G_{nm}) \leq \lambda(E_n) + 1/n$ and take $H_n = \bigcap_{m=1}^{\infty} G_{nm}$.)

Write $V_m = \bigcap_{k=m}^{\infty} H_k$ and $V = \bigcup_{m=1}^{\infty} V_m$. These sets are all measurable because we choose the $\{H_k\}$ to be measurable.

We obtain

$$\lambda(V) = \lim_{m \rightarrow \infty} \lambda(V_m).$$

But $E_m \subset V_m \subset H_m$ so that $V \supset E$ and $\lambda(E_m) = \lambda(V_m) = \lambda(H_m)$. Consequently

$$\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \lambda(V) = \lim_{m \rightarrow \infty} \lambda(V_m) = \lim_{m \rightarrow \infty} \lambda(E_m).$$

This completes the proof. ■

7.6.4 Existence of nonmeasurable sets

We turn now to a search for Lebesgue nonmeasurable sets. The first proof that nonmeasurable sets must exist is due to G. Vitali (1875–1932). It uses the axiom of choice which has to this point not been needed in the text.

Theorem 7.15 *There exist subsets of $\mathbb{R}$ that are not Lebesgue measurable.*

Proof. Let $I = [-\frac{1}{2}, \frac{1}{2}]$. We define an equivalence relation on this interval by relating points to rational numbers; we use Q to denote the set of all rationals. For $x, y \in I$ write $x \sim y$ if $x - y \in Q$. For all $x \in I$, let

$$K(x) = \{y \in I : x - y \in Q\} = \{x + r \in I : r \in Q\}.$$

We show that $\sim$ is an equivalence relation. It is clear that $x \sim x$ for all $x \in I$ and that if $x \sim y$ then $y \sim x$. To show transitivity of $\sim$, suppose that $x, y, z \in I$ and $x - y = r_1$ and $y - z = r_2$ for $r_1, r_2 \in Q$. Then $x - z = (x - y) + (y - z) = r_1 + r_2$, so $x \sim z$. Thus the set of all equivalence classes $K(x)$ forms a partition of I : $\bigcup_{x \in I} K(x) = I$, and if $K(x) \neq K(y)$, then $K(x) \cap K(y) = \emptyset$.

Let A be a set containing exactly one member of each equivalence class. (The existence of such a set A follows from the axiom of choice.) We show that A is nonmeasurable. Let $0 = r_0, r_1, r_2, \dots$ be an enumeration of $Q \cap [-1, 1]$, and define

$$A_k = \{x + r_k : x \in A\}$$

so that A_k is obtained from A by the translation $x \rightarrow x + r_k$.

Then

$$[-\frac{1}{2}, \frac{1}{2}] \subset \bigcup_{k=0}^{\infty} A_k \subset [-\frac{3}{2}, \frac{3}{2}]. \tag{7.4}$$

To verify the first inclusion, let $x \in [-\frac{1}{2}, \frac{1}{2}]$ and let x_0 be the representative of $K(x)$ in A . We have $\{x_0\} = A \cap K(x)$. Then $x - x_0 \in Q \cap [-1, 1]$, so there exists k such that $x - x_0 = r_k$. Thus $x \in A_k$. The second inclusion is immediate: the set A_k is the translation of $A \subset [-\frac{1}{2}, \frac{1}{2}]$ by the rational number $r_k \in [-1, 1]$.

Suppose now that A is measurable. It is easy to see that then each of the translated sets A_k is also measurable and that $\lambda(A_k) = \lambda(A)$ for every k . But the sets $\{A_i\}$ are pairwise disjoint. If $z \in A_i \cap A_j$ for $i \neq j$, then $x_i = z - r_i$ and $x_j = z - r_j$ are in different equivalence classes. This is impossible, since $x_i - x_j \in Q$. It now follows from (7.4) and the countable additivity of λ for measurable set that

$$1 = \lambda\left(-\frac{1}{2}, \frac{1}{2}\right] \leq \lambda\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \lambda(A_k) \leq \lambda\left(-\frac{3}{2}, \frac{3}{2}\right] = 3. \quad (7.5)$$

Let $\alpha = \lambda(A) = \lambda(A_k)$. From (7.5), we infer that

$$1 \leq \alpha + \alpha + \cdots \leq 3. \quad (7.6)$$

But it is clear that no number α can satisfy both inequalities in (7.6). The first inequality implies that $\alpha > 0$, but the second implies that $\alpha = 0$. Thus A is nonmeasurable. ■

The proof has invoked the axiom of choice in order to construct the nonmeasurable set. One might ask whether it is possible to give a more constructive proof, one that does not use this principle. This question belongs to the subject of logic rather than analysis, and the logicians have answered it. In 1964, R. M. Solovay showed that, in Zermelo–Fraenkel set theory with a weaker assumption than the axiom of choice, it is consistent that all sets are Lebesgue measurable. On the other hand, the existence of nonmeasurable sets does not imply the axiom of choice. Thus it is no accident that our proof had to rely on the axiom of choice: it would have to appeal to some further logical principle in any case.²

²See also K. Ciesielski, “How good is Lebesgue measure?” *Math. Intelligencer* 11(2), 1989, pp. 54–58, for a discussion of material related to this section and for references to the literature. That same author’s text, *Set Theory for the Working Mathematician*, Cambridge University Press, London (1997) is an excellent source for students wishing to go deeper into these ideas.

7.7 Measurable functions

Definition 7.16 An arbitrary function $f : \mathbb{R} \rightarrow \mathbb{R}$ is measurable if for any real number r

$$A_r = \{x \in \mathbb{R} : f(x) < r\}$$

is a measurable set.

A function $f : [a, b] \rightarrow \mathbb{R}$ would be measurable if there is a measurable function $g : \mathbb{R} \rightarrow \mathbb{R}$ and $f(x) = g(x)$ for all $x \in [a, b]$.

Exercises

Exercise 666 Let f be a measurable function. Show that each of $|f|$, $[f]^+$, and $[f]^-$ must also be measurable. □

Exercise 667 Show that the function $f(x) = \chi_A(x)$ is measurable if and only if the set A is a measurable set. □

7.7.1 Continuous functions are measurable

Lemma 7.17 A function $f : \mathbb{R} \rightarrow \mathbb{R}$ that is continuous everywhere is measurable.

Proof. To prove that f is measurable we need to verify that, for any real number r ,

$$A_r = \{x \in \mathbb{R} : f(x) < r\}$$

is a measurable set. But we already know that, for continuous functions, such sets are open. ■

We know too that a continuous function $f : [a, b] \rightarrow \mathbb{R}$ is also measurable by our definition since f agrees on $[a, b]$ with the continuous function g defined by $g(t) = f(t)$ for $a \leq t \leq b$, $g(t) = g(b)$ for $t > b$, and $g(t) = g(a)$ for $t < a$.

7.7.2 Derivatives and integrable functions are measurable

Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ is almost everywhere the derivative of some function. Then f is measurable³. If we combine that fact with the definition of the calculus integral we see that all integrable functions must be measurable.

Lemma 7.18 *A function $f : \mathbb{R} \rightarrow \mathbb{R}$ that is almost everywhere the derivative of some function is measurable.*

Proof. We suppose that $F : \mathbb{R} \rightarrow \mathbb{R}$ and $F'(x) = f(x)$ almost everywhere, say everywhere in $\mathbb{R} \setminus N$ where N is a set of measure zero. Consider the set $E = \{x : \overline{DF}(x) > r\}$ for any r . Let m, n be positive integers and define β_{mn} to be the covering relation consisting of all pairs $([u, v], w)$ for which $u \leq w \leq v$, and for which $0 < v - u < 1/m$ and

$$\frac{F(v) - F(u)}{v - u} \geq r + 1/n.$$

Write

$$E_{mn} = \bigcup \{[u, v] : ([u, v], w) \in \beta_{mn}\}.$$

Each set E_{mn} is thus a fairly simple object: it is a union of a family of compact intervals. In Lemma 5.16 we have seen that this means it has a simple structure: it differs from an open set by a countable set. In particular each E_{mn} is an measurable set. We check that

$$E = \bigcup_{n=1}^{\infty} \bigcap_{m=1}^{\infty} E_{mn}. \quad (7.7)$$

To begin suppose that $x \in E$. Then $\overline{DF}(x) > r$. There must be at least one integer n with $\overline{DF}(x) > r + 1/n$. Moreover, for every integer m there would have to be at least one compact interval $[u, v]$ containing x with length less than $1/m$ so that

$$\frac{F(v) - F(u)}{v - u} \geq r + 1/n.$$

Hence x is a point in the set on the right-hand side of the proposed identity. Conversely, should x belong to that set, then there is at least one n so that for all m , x belongs to E_{mn} . It would follow that $\overline{DF}(x) > r$ and so $x \in E$.

The identity (7.7) now exhibits E as a combination of sequences of measurable sets and so E too is an measurable

³A theorem of Lusin states the converse: if f is measurable then there is a continuous function F for which $F'(x) = f(x)$ almost everywhere. This should not be confused with the fundamental theorem of the calculus.

set because the measurable sets form a Borel family (Theorem 7.13). Finally then

$$\{x : f(x) > r\} = (\{x : \overline{D}F(x) > r\} \cap [\mathbb{R} \setminus N]) \cup N'$$

where N' is an appropriate subset of N . This exhibits the set $\{x : f(x) > r\}$ as the union of a measurable set and a set of measure zero. Consequently that set is measurable. This is true for all r and verifies that f is a measurable function. ■

Corollary 7.19 *If $f : [a, b] \rightarrow \mathbb{R}$ is integrable then f is measurable.*

Exercise 668 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$. Show that the set of points where f is differentiable is a measurable set.*

Answer □

7.7.3 Simple functions

A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is *simple* if there is a finite collection of measurable sets $E_1, E_2, E_3, \dots, E_n$ and real numbers $r_1, r_2, r_3, \dots, r_n$ so that

$$f(x) = \sum_{k=1}^n r_k \chi_{E_k}(x)$$

for all real x .

Lemma 7.20 *Any simple function is measurable.*

Proof. Suppose that

$$f(x) = \sum_{k=1}^n r_k \chi_{E_k}(x)$$

and s is any real number. It is easy to sort out, for any value of s , exactly what the set

$$A_s = \{x : f(x) < s\}$$

must be in terms of the sets $\{E_k\}$. In each case we see that A_s is some simple combination of measurable sets and so is itself measurable. ■

7.7.4 Series of simple functions

Theorem 7.21 Every nonnegative, measurable function $f : \mathbb{R} \rightarrow \mathbb{R}$ can be written as the sum of a series of nonnegative simple functions by the following inductive procedure: Take $\{r_k\}$ to be any sequence of positive numbers for which $r_k \rightarrow 0$ and $\sum_{k=1}^{\infty} r_k = +\infty$. Define the sets

$$A_k = \left\{ x : f(x) \geq r_k + \sum_{j < k} r_j \chi_{A_j}(x) \right\}$$

inductively, starting with $A_0 = \emptyset$. Then

$$f(x) = \sum_{k=1}^{\infty} r_k \chi_{A_k}(x)$$

at every x .

The proof is just a matter of deciding whether and why this works.

Exercises

Exercise 669 Prove Theorem 7.21.

□

Exercise 670 Show that the following procedure expresses a nonnegative, measurable function $f : \mathbb{R} \rightarrow \mathbb{R}$ as a nondecreasing limit of a sequence $\{f_k\}$ of simple functions: Fix an integer k . Subdivide $[0, k]$ into subintervals

$$[(j-1)2^{-k}, j2^{-k}] \quad (j = 1, 2, 3, \dots, k2^k)$$

and, for all $x \in [a, b]$, define $f_k(x)$ to be $(j-1)2^{-k}$ if

$$(j-1)2^{-k} \leq f(x) < j2^{-k}$$

and to be k if $f(x) \geq k$.

□

Exercise 671 In the preceding exercise show that, if f is bounded, then f is the uniform limit of the sequence of simple functions $\{f_k\}$.

□

7.7.5 Limits of measurable functions

Theorem 7.22 Let $f_n : \mathbb{R} \rightarrow \mathbb{R}$ be a sequence of measurable functions. Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ is a function for which

$$f(x) = \lim_{n \rightarrow \infty} f_n(x)$$

for almost every x . Then f is measurable.

Proof. We fix a real number r and verify that

$$\{x \in \mathbb{R} : f(x) < r\}$$

is a measurable set. We use the fact that sets of the form

$$\{x \in \mathbb{R} : f_n(x) < s\}$$

are measurable. This follows from the measurability of each function f_n .

Let N be the null set consisting of points x where we do not have

$$f(x) = \lim_{n \rightarrow \infty} f_n(x)$$

and let $E = \mathbb{R} \setminus N$. Then both E and N are measurable.

We claim the following set identity:

$$\{x \in E : f(x) < r\} = \bigcup_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \{x \in E : f_n(x) < r - 1/k\}.$$

This is a matter of close interpretation. If x_0 belongs to the simple set on the left of the proposed identity, then $x_0 \in E$ and $f(x_0) < r$. There must exist a k so that $f(x_0) < r - 1/k$. Then there must exist an integer m so that $f_n(x) < r - 1/k$ for all $n \geq m$. That places x_0 in the set on the right.

In the other direction if x_0 belongs to the complicated set on the right of the proposed identity, then for some k and m , $f_n(x_0) < r - 1/k$ for all $n \geq m$. It follows that $f(x_0) \leq r - 1/k < r$. That places x_0 in the set on the left.

Each set

$$\{x \in E : f_n(x) < r - 1/k\} = E \cap \{x \in \mathbb{R} : f_n(x) < r - 1/k\}$$

thus is measurable since it is the intersection of a measurable set and an open set. As measurable sets form a Borel family the intersections and unions of these sets remain measurable.

Finally then

$$\{x \in \mathbb{R} : f(x) < r\}$$

is seen to be the union of the measurable set

$$\{x \in E : f(x) < r\}$$

and some subset of N . This checks the measurability of the function f . ■

7.8 Construction of the integral

We now give Lebesgue's construction of the integral in a series of steps, starting with characteristic functions, then simple functions, then nonnegative measurable functions, and finally all absolutely integrable functions.

7.8.1 Characteristic functions of measurable sets

Lemma 7.23 *Let E be a subset of an interval $[a, b]$. Then χ_E is integrable on $[a, b]$ if and only if E is a measurable set, and in that case*

$$\lambda(E) = \int_a^b \chi_E(x) dx.$$

Proof. For any set $E \subset [a, b]$, measurable or not, we can easily establish the (Exercise 661) identity

$$\lambda^*(E) = \int_a^b \overline{\chi_E}(x) dx.$$

The two concepts in this identity are defined by the same process. Thus the proof of the lemma depends only on showing that integrability of $\chi_E(x)$ is equivalent to the measurability of E .

We already know that if $\chi_E(x)$ is integrable then it is a measurable function. But this can happen only if E is a measurable set. Conversely let us suppose that E is measurable and verify that χ_E is integrable on $[a, b]$. In fact we show that this function satisfies the McShane criterion on this interval (see Exercise 657).

Since E is measurable we know that

$$\lambda(E) + \lambda([a, b] \setminus E) = b - a.$$

Let $\varepsilon > 0$. Select open sets $E \subset G_1$ and $[a, b] \setminus G_2$ so that

$$\lambda(G_1) < \lambda(E) + \varepsilon/2$$

and

$$\lambda(G_2) < \lambda([a, b] \setminus E) + \varepsilon/2.$$

Then, use the identity

$$\lambda(G_1 \cup G_2) = \lambda(G_1) + \lambda(G_2) - \lambda(G_1 \cap G_2)$$

to get

$$\begin{aligned} \lambda(G_1 \cap G_2) &= \lambda(G_1) + \lambda(G_2) - \lambda(G_1 \cup G_2) \\ &< [\lambda(E) + \varepsilon/2] + [\lambda([a, b] \setminus E) + \varepsilon/2] - (b - a) = \varepsilon. \end{aligned}$$

This will enable us to apply the McShane criterion to establish that χ_E is integrable on $[a, b]$. Define β as the collection of all pairs $([u, v], w)$ for which either $w \in E$ and $[u, v] \subset G_1$ or $w \in [a, b] \setminus E$ and $[u, v] \subset G_2$. This is a full cover of $[a, b]$. Choose any two partitions π, π' of $[a, b]$ contained in β . We compute

$$\sum_{([u, v], w) \in \pi} \sum_{([u', v'], w') \in \pi'} |\chi_E(w) - \chi_E(w')| \lambda([u, v] \cap [u', v']). \quad (7.8)$$

Note, in this sum, that terms for which both w and w' are in E or for which neither is in E vanish. Terms for which $w \in E$ and $w' \in [a, b] \setminus E$ must have $|\chi_E(w) - \chi_E(w')| = 1$, $[u, v] \subset G_1$ and $[u', v'] \subset G_2$. In particular $[u, v] \cap [u', v'] \subset (G_1 \cap G_2)$. The same is true if $w' \in E$ and $w \in [a, b] \setminus E$. Remembering that $\lambda(G_1 \cap G_2) < \varepsilon$, we see that the sum in (7.8) is smaller than ε . By the McShane criterion χ_E is integrable on $[a, b]$. ■

7.8.2 Characterizations of measurable sets

As corollaries we obtain a number of characterizations of measurable sets, including the original Lebesgue definition which is assertion (3). Assertion (4) is known as Carathéodory's criterion.

Corollary 7.24 *Let E be a set of real numbers. Then the following assertions are equivalent:*

1. E is measurable.
2. χ_E is integrable on every compact interval $[a, b]$.
3. For every compact interval $[a, b]$,

$$\lambda([a, b] \cap E) + \lambda([a, b] \setminus E) = b - a. \quad (7.9)$$

4. For every set $T \subset \mathbb{R}$,

$$\lambda(T) \geq \lambda(T \cap E) + \lambda(T \setminus E). \quad (7.10)$$

5. For every $\varepsilon > 0$ and every compact interval $[a, b]$, there is a full cover β of $[a, b]$ so that

$$\sum_{([u, v], w) \in \pi} \sum_{([u', v'], w') \in \pi'} \lambda([u, v] \cap [u', v']) < \varepsilon$$

whenever π, π' are subpartitions of $[a, b]$ with $\pi \subset \beta[E]$ and $\pi' \subset \beta[[a, b] \setminus E]$.

Proof. First note that a set E is measurable if and only if $E \cap [a, b]$ is measurable for every compact interval $[a, b]$. In one direction this is because $[a, b]$ is a measurable set (it is closed) and the intersection of measurable sets is also measurable. In the other direction, if $E \cap [a, b]$ is measurable for every compact interval $[a, b]$, then $E = \bigcup_{n=1}^{\infty} E \cap [-n, n]$ expresses E as a measurable set.

The first three conditions (a), (b), and (c) we have explicitly shown to be equivalent in the proof of the lemma. Let us check that (d) implies (c). Observe that the inequality,

$$\lambda(T) \leq \lambda(T \cap E) + \lambda(T \setminus E)$$

holds in general, so that the condition (7.10) is equivalent to the assertion of equality:

$$\lambda(T) = \lambda(T \cap E) + \lambda(T \setminus E).$$

Thus (c) is a special case of (d) with $T = [a, b]$. On the other hand, (a) implies (d). Measurability of E implies that E and

$\mathbb{R} \setminus E$ are disjoint measurable sets for which

$$\lambda(T) = \lambda(T \cap E) + \lambda(T \setminus E)$$

must hold for any set $T \subset \mathbb{R}$. Finally the fifth condition (e) is just a rewriting of the McShane criterion for integrability of the function χ_E on $[a, b]$. We have seen in the proof of the lemma that measurability of $E \cap [a, b]$ is equivalent to that criterion applied to χ_E on $[a, b]$. ■

7.8.3 Integral of simple functions

Recall that a function $f : \mathbb{R} \rightarrow \mathbb{R}$ is *simple* if there is a finite collection of measurable sets $E_1, E_2, E_3, \dots, E_n$ and real numbers $r_1, r_2, r_3, \dots, r_n$ so that

$$f(x) = \sum_{k=1}^n r_k \chi_{E_k}(x)$$

for all real x . Since this is a finite linear combination it follows from the integration theory and the integration of characteristic functions (Lemma 7.23) that such a function is necessarily integrable on any compact interval $[a, b]$ and that

$$\int_a^b f(x) dx = \sum_{k=1}^n \left(\int_a^b r_k \chi_{E_k}(x) dx \right) = \sum_{k=1}^n r_k \lambda(E_k \cap [a, b]).$$

Thus the integral of simple functions can be constructed from the values of the function in a finite number of steps using the Lebesgue measure.

7.8.4 Integral of nonnegative measurable functions

We have seen (Theorem 7.21) that every nonnegative measurable function can be represented by simple functions. Consequently the integral of such a function can be constructed.

Theorem 7.25 Let f be a nonnegative, measurable function on an interval $[a, b]$. Then, for any representation of f as the sum of a series of nonnegative, simple functions

$$f(x) = \sum_{k=1}^{\infty} f_n(x) \quad (a \leq x \leq b)$$

the identity

$$\int_a^b f(x) dx = \sum_{k=1}^{\infty} \left(\int_a^b f_n(x) dx \right)$$

must hold (finite or infinite). Moreover f is integrable on $[a, b]$ if and only if this series of integrals converges to a finite value.

Proof. This requires only an appeal to the monotone convergence theorem. ■

Corollary 7.26 Let f be a nonnegative, measurable function on an interval $[a, b]$. Then

$$\int_a^b f(x) dx$$

exists (finitely or infinitely). Moreover f is integrable on $[a, b]$ if and only if this value is finite.

Proof. This follows from the theorem. ■

7.8.5 Fatou's Lemma

Theorem 7.27 (Fatou's lemma) Let f_n be a sequence of nonnegative, measurable functions defined at every point of an interval $[a, b]$. Then, assuming that

$$f(x) = \liminf_{n \rightarrow \infty} f_n(x)$$

is finite almost everywhere,

$$\int_a^b \liminf_{n \rightarrow \infty} f_n(x) dx \leq \liminf_{n \rightarrow \infty} \int_a^b f_n(x) dx..$$

Proof. Fatou's lemma is proved using the monotone convergence theorem, Theorem 3.43. Let f denote the limit inferior of the f_n . For every natural number k define the function

$$g_k(x) = \inf_{n \geq k} f_n(x).$$

Then the sequence $g_1, g_2, \dots$ is a nondecreasing sequence of measurable functions and converges pointwise to f . For $k \leq n$, we have $g_k(x) \leq f_n(x)$, so that

$$\int_a^b g_k(x) dx \leq \int_a^b f_n(x) dx,$$

hence

$$\int_a^b g_k(x) dx \leq \inf_{n \geq k} \int_a^b f_n(x) dx.$$

Using the monotone convergence theorem, the last inequality, and the definition of the limit inferior, it follows that

$$\int_a^b \liminf_{n \rightarrow \infty} f_n(x) dx = \lim_{k \rightarrow \infty} \int_a^b g_k(x) dx \leq \lim_{k \rightarrow \infty} \inf_{n \geq k} \int_a^b f_n(x) dx = \liminf_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

■

Exercises

Exercise 672 On the interval $[0, 1]$ for every natural number n define

$$f_n(x) = \begin{cases} n & \text{for } x \in (0, 1/n), \\ 0 & \text{otherwise.} \end{cases}$$

Show that

$$\int_0^1 \liminf_{n \rightarrow \infty} f_n(x) dx < \liminf_{n \rightarrow \infty} \int_0^1 f_n(x) dx.$$

□

Exercise 673 On the interval $[0, \infty)$ for every natural number n define

$$f_n(x) = \begin{cases} \frac{1}{n} & \text{for } x \in [0, n], \\ 0 & \text{otherwise.} \end{cases}$$

Show that $\{f_n\}$ is uniformly convergent and that

$$\int_0^\infty \liminf_{n \rightarrow \infty} f_n(x) dx < \liminf_{n \rightarrow \infty} \int_0^\infty f_n(x) dx.$$

□

Exercise 674 On the interval $[0, \infty)$ for every natural number n define

$$f_n(x) = \begin{cases} -\frac{1}{n} & \text{for } x \in [n, 2n], \\ 0 & \text{otherwise.} \end{cases}$$

Show that $\{f_n\}$ is uniformly convergent and that the inequality in Fatou's lemma

$$\int_0^\infty \liminf_{n \rightarrow \infty} f_n(x) dx \leq \liminf_{n \rightarrow \infty} \int_0^\infty f_n(x) dx.$$

fails.

□

Exercise 675 (reverse Fatou lemma) Let $\{f_n\}$ be a sequence of measurable functions defined on an interval $[a, b]$. Suppose that there exists a Lebesgue integrable function g on $[a, b]$ such that $f_n \leq g$ for all n . Show that

$$\int_a^b \limsup_{n \rightarrow \infty} f_n(x) dx \geq \limsup_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

Answer □

Exercise 676 (dominated convergence theorem) Let $\{f_n\}$ be a sequence of measurable functions defined on an interval $[a, b]$. Assume that the sequence converges pointwise and is dominated by some nonnegative, Lebesgue integrable function g . Then the pointwise limit is an integrable function and

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \lim_{n \rightarrow \infty} f_n(x) dx.$$

To say that the sequence is "dominated" by g means that $|f_n(x)| \leq g(x)$ for all natural numbers n and all points x in $[a, b]$.

Answer □

7.8.6 Derivatives of functions of bounded variation

As a consequence of Lebesgue's program to this point we can prove some facts about derivatives of monotonic functions and derivatives of functions of bounded variation. These are due to Lebesgue, but our proofs are rather easier since we do not need much of the measure theory to obtain them.

Theorem 7.28 *Let $F : [a, b] \rightarrow \mathbb{R}$ be a function of bounded variation. Then $F'(x)$ exists almost everywhere in $[a, b]$ and*

$$\int_a^b |F'(x)| dx \leq V(F, [a, b]).$$

Proof. We know from the Lebesgue differentiation theorem that F is a.e. differentiable. Let $f(x) = |F'(x)|$ at every point at which $F'(x)$ exists and as zero elsewhere. Then f is a nonnegative function. At every point w in $[a, b]$ there is a $\delta > 0$ so that, whenever $u \leq w \leq v$ and $0 < v - u < \delta$,

$$f(w) - \varepsilon \leq \frac{|F(v) - F(u)|}{v - u}.$$

At points w where $f(w) = 0$ this is obvious, while at points w where $F'(w)$ exists this follows from the definition of the derivative.

Take β as the collection of all pairs $([u, v], w)$ subject to the requirement only that

$$|F(v) - F(u)| > [f(w) - \varepsilon](v - u)$$

if $w \in [a, b]$ and $[u, v] \subset [a, b]$. This collection β is a full cover.

Every partition $\pi \subset \beta$ of the interval $[a, b]$ satisfies

$$\sum_{([u, v], w) \in \pi} [f(w) - \varepsilon](v - u) < \sum_{([u, v], w) \in \pi} |F(v) - F(u)| \leq V(F, [a, b]).$$

It follows that

$$-\varepsilon(b - a) + \int_a^b f(x) dx \leq V(F, [a, b]).$$

Since ε is an arbitrary positive number,

$$\int_a^b f(x) dx \leq V(F, [a, b]).$$

Since f is almost everywhere a derivative it is necessarily measurable. Thus we may use the integral in place of the upper integral. ■

Corollary 7.29 *Let $F : [a, b] \rightarrow \mathbb{R}$ be a nondecreasing function. Then $F'(x)$ exists almost everywhere in $[a, b]$ and*

$$\int_a^b F'(x) dx \leq F(b) - F(a).$$

Corollary 7.30 (Lebesgue decomposition) *Let $F : [a, b] \rightarrow \mathbb{R}$ be a continuous, nondecreasing function. Then $F'(x)$ exists almost everywhere in $[a, b]$ and*

$$F(t) = \int_a^t F'(x) dx + S(t) \quad (a \leq t \leq b)$$

expresses F as the sum of an integral and a continuous, nondecreasing singular function.

Proof. Simply define

$$S(t) = F(t) - \int_a^t F'(x) dx \quad (a \leq t \leq b).$$

Check that $S'(t) = 0$ almost everywhere (trivial) and so S is singular. That S is continuous is evident since it is the difference of two continuous functions. That S is nondecreasing follows from the theorem, since

$$S(d) - S(c) = F(d) - F(c) - \int_c^d F'(x) dx \geq 0$$

for any $[c, d] \subset [a, b]$. ■

7.8.7 Characterization of the Lebesgue integral

Recall that a function f is Lebesgue integrable on an interval $[a, b]$ if both f and $|f|$ are integrable on that interval.

Theorem 7.31 *Let $f : [a, b] \rightarrow \mathbb{R}$. Then f is Lebesgue integrable if and only if f is measurable and*

$$\int_a^b |f(x)| dx < \infty.$$

Proof. We know, from Exercise 666, that the functions $|f|$, $[f]^+$, and $[f]^-$ are also measurable. The finiteness of this integral implies (by Corollary 7.26) that each of these functions are integrable. In particular both functions $f = [f]^+ - [f]^-$ and $|f|$ are integrable. Thus f must be absolutely integrable. Conversely if f is absolutely integrable, this means that $|f|$ is integrable and consequently, by definition, it has a finite integral. ■

Our final theorem for Lebesgue's program shows that the integral is constructible by his methods for all Lebesgue integrable functions. We see in the next section that this is as far as one can go.

Theorem 7.32 *If f is Lebesgue integrable on a compact interval $[a, b]$ then f , $|f|$, $[f]^+$, and $[f]^-$ are measurable and*

$$\int_a^b |f(x)| dx = \int_a^b [f(x)]^+ dx + \int_a^b [f(x)]^- dx$$

and

$$\int_a^b f(x) dx = \int_a^b [f(x)]^+ dx - \int_a^b [f(x)]^- dx$$

Proof. If f is Lebesgue integrable then we know that f and $|f|$ are integrable. It follows that $[f]^+ = (f + |f|)/2$ and $[f]^- = (|f| - f)/2$ are both integrable. All functions are measurable since all are integrable. Since

$$|f(x)| = [f(x)]^+ + [f(x)]^-$$

and

$$f(x) = [f(x)]^+ - [f(x)]^-$$

the integration formulas are immediately available. ■

7.8.8 McShane's Criterion

Lebesgue's integral can also be characterized by the McShane criterion. Using normal inequality techniques we easily observe that the expression

$$\left| \sum_{(I, w) \in \pi} \sum_{(I', w') \in \pi'} [f(w) - f(w')] \lambda(I \cap I') \right| < \varepsilon \quad (7.11)$$

that we use for the second Cauchy criterion must be smaller than a quite similar expression:

$$\left| \sum_{(I,w) \in \pi} \sum_{(I',w') \in \pi'} [f(w) - f(w')] \lambda(I \cap I') \right| \leq \sum_{(I,w) \in \pi} \sum_{(I',w') \in \pi'} |f(w) - f(w')| \lambda(I \cap I').$$

It takes a sharp (and young) eye to spot the difference, but the larger side of this inequality may be strictly larger. This leads to a stronger integrability criterion than that in the second Cauchy criterion. This is the motivation for the criterion, named after E. J. McShane. We prove that McShane's criterion is a necessary and sufficient condition for Lebesgue integrability.

Definition 7.33 (McShane's criterion) A function $f : [a, b] \rightarrow \mathbb{R}$ is said to satisfy McShane's criterion on $[a, b]$ provided that for all $\varepsilon > 0$ a full cover β can be found so that

$$\sum_{(I,w) \in \pi} \sum_{(I',w') \in \pi'} |f(w) - f(w')| \lambda(I \cap I') < \varepsilon$$

for all partitions π, π' of $[a, b]$ contained in β .

Theorem 7.34 If f satisfies McShane's criterion on $[a, b]$ then f is absolutely integrable, i.e., both f and $|f|$ are integrable there and

$$-\int_a^b f(x) dx \leq \int_a^b |f(x)| dx \leq \int_a^b f(x) dx.$$

Proof.

Theorem 7.35 Let $f : [a, b] \rightarrow \mathbb{R}$. Then f is Lebesgue integrable on an interval if and only if it satisfies McShane's criterion on that interval.

Proof. It is immediate that if f satisfies McShane's criterion it also satisfies Cauchy's second criterion. Thus the function f is integrable. We then observe that, since

$$||f(x)| - |f(x')|| \leq |f(x) - f(x')|,$$

it is clear that whenever f satisfies McShane's criterion so too does $|f|$. Thus $|f|$ too is integrable on $[a, b]$. The inequalities of the theorem simply follow from the inequalities $-|f(x)| \leq f(x) \leq |f(x)|$ which hold for all x .

Here is the proof in the other direction. To simplify the notation let us write

$$S(f, \pi, \pi') = \sum_{([u,v], w) \in \pi} \sum_{([u',v'], w') \in \pi'} |f(w) - f(w')| \lambda([u, v] \cap [u', v']) \quad (7.12)$$

for any two partitions π, π' of $[a, b]$. Some preliminary computations will help. If $g_1, g_2, \dots, g_n$ are functions on $[a, b]$ then,

$$S\left(\sum_{i=1}^n g_i, \pi, \pi'\right) \leq \sum_{i=1}^n S(g_i, \pi, \pi'). \quad (7.13)$$

If

$$\overline{\int_a^b |f(x)| dx} < t$$

then there must exist a full cover β with the property that for any two partitions π, π' of $[a, b]$ from β ,

$$S(f, \pi, \pi') < 2t. \quad (7.14)$$

Finally

$$S(f, \pi, \pi') \leq \sup\{|f(t)| : a \leq t \leq b\} \cdot 2(b-a). \quad (7.15)$$

Each of the statements (7.13), (7.14), and (7.15) require only simple computations that we leave to the reader.

Now for our argument. We assume that f is absolutely integrable and verify the criterion. But f can be written as a difference of two nonnegative integrable functions. If both of these satisfy the criterion then, using (7.13) we deduce that so too does f . Consequently for the remainder of the proof we assume that f is nonnegative and integrable.

The first step is to observe that every characteristic function of a measurable set satisfies the McShane criterion. This is proved in Lemma 7.23. Using (7.13) we easily deduce, as our second step, that every nonnegative simple function also satisfies the McShane criterion.

The third step is to show that every nonnegative, bounded measurable function also satisfies this criterion. But such a function is the uniform limit of a sequence of nonnegative simple functions. It follows then, from (7.15), that such functions satisfy the McShane criterion. For if f is a bounded measurable function, $\varepsilon > 0$, choose a simple function g so that

$$|f(t) - g(t)| < \varepsilon / (4[b-a])$$

7.8. CONSTRUCTION OF THE INTEGRAL

297

for all $a \leq t \leq b$. Now using McShane's criterion on g we can select a full cover β for which $S(g, \pi, \pi') < \varepsilon/2$ for all partitions π, π' of $[a, b]$ from β . Then

$$S(f, \pi, \pi') \leq S(f - g, \pi, \pi') + S(g, \pi, \pi') \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

The final step requires an appeal to the monotone convergence theorem. Set $f_N(t) = \min\{N, f(t)\}$ and use the monotone convergence theorem to find an integer N large enough so that

$$\int_a^b [f(x) - f_N(x)] dx < \varepsilon/4.$$

Using (7.14) select a full cover β_1 for which $S(f - f_N, \pi, \pi') < \varepsilon/2$ for all partitions π, π' of $[a, b]$ from β_1 . Select a full cover β_2 for which $S(f_N, \pi, \pi') < \varepsilon/2$ for all partitions π, π' of $[a, b]$ from β_2 . Then set $\beta = \beta_1 \cap \beta_2$. This is a full cover and we can check that

$$S(f, \pi, \pi') \leq S(f - f_N, \pi, \pi') + S(f_N, \pi, \pi') \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

for all partitions π, π' of $[a, b]$ from β . This verifies the McShane criterion for an arbitrary nonnegative integrable function f . ■

Exercises

Exercise 677 Suppose that each of the functions $f_1, f_2, \dots, f_n : [a, b] \rightarrow \mathbb{R}$ satisfies McShane's criterion on a compact interval $[a, b]$ and that a function $L : \mathbb{R}^n \rightarrow \mathbb{R}$ is given satisfying

$$|L(x_1, x_2, \dots, x_n) - L(y_1, y_2, \dots, y_n)| \leq M \sum_{i=1}^n |x_i - y_i|$$

for some number M and all $(x_1, x_2, \dots, x_n)$ and $(y_1, y_2, \dots, y_n)$ in $\mathbb{R}^n$. Show that the function $g(x) = L(f_1(x), f_2(x), \dots, f_n(x))$ satisfies McShane's criterion on $[a, b]$. □

Exercise 678 Let $F, f : \mathbb{R} \rightarrow \mathbb{R}$. A necessary and sufficient condition in order that f be the derivative of F at each point is that for every $\varepsilon > 0$ there is a full cover β of the real line with the property that for every compact interval $[a, b]$ and every partition $\pi \subset \beta$ of $[a, b]$,

$$\sum_{(I, x) \in \pi} |\Delta F(I) - f(x)\lambda(I)| < \varepsilon\lambda([a, b]). \quad (7.16)$$

Answer $\square$

Exercise 679 (Freiling's criterion) Let $f : \mathbb{R} \rightarrow \mathbb{R}$. Show that necessary and sufficient condition⁴ in order that f be the derivative of some function F at each point is that for every $\varepsilon > 0$ there is a full cover β of the real line with the property that for every compact interval $[a, b]$ and every pair of partitions $\pi_1, \pi_2 \subset \beta$ of $[a, b]$,

$$\left| \sum_{(I,z) \in \pi} \sum_{(I',z') \in \pi'} [f(z) - f(z')] \lambda(I \cap I') \right| < \varepsilon \lambda([a, b]). \quad (7.17)$$

Answer $\square$

Exercise 680 Let $f : \mathbb{R} \rightarrow \mathbb{R}$. Characterize the following property: for every $\varepsilon > 0$ there is a full cover β of the real line with the property that for every compact interval $[a, b]$ and every pair of partitions $\pi_1, \pi_2 \subset \beta$ of $[a, b]$,

$$\sum_{(I,z) \in \pi} \sum_{(I',z') \in \pi'} |f(z) - f(z')| \lambda(I \cap I') < \varepsilon \lambda([a, b]).$$

 $\square$

7.8.9 Nonabsolutely integrable functions

A function f is *nonabsolutely integrable* on an interval $[a, b]$ if it is integrable, but not absolutely integrable there, i.e., f is integrable $[a, b]$ but $|f|$ is not integrable. Lebesgue's program will not construct the integral of a nonabsolutely integrable function. The only method that his program offers is the hope that

$$\int_a^b f(x) dx = \int_a^b [f(x)]^+ dx - \int_a^b [f(x)]^- dx?$$

Theorem 7.36 If f is nonabsolutely integrable on a compact interval $[a, b]$ then

$$\int_a^b |f(x)| dx = \int_a^b [f(x)]^+ dx = \int_a^b [f(x)]^- dx = \infty.$$

⁴This is from Chris Freiling, *On the problem of characterizing derivatives*. Real Anal. Exchange 23 (1997/98), no. 2, 805–812.

Proof. If f is nonabsolutely integrable then it is measurable. It follows from Exercise 666 that the functions $|f|$, $[f]^+$, and $[f]^-$ are also measurable. If, for example,

$$\int_a^b [f(x)]^+ dx < \infty,$$

contrary to what we wish to prove, then we must conclude (from Theorem 7.31) that $[f]^+$ is integrable. But if $[f]^+$ is integrable then from the identity

$$[f(x)]^- = [f(x)]^+ - f(x)$$

we could conclude that $[f]^-$ must also be integrable and consequently each of the functions f , $|f|$, $[f]^+$, and $[f]^-$ must be integrable, contradicting the hypothesis of the theorem. ■

7.9 The Lebesgue integral as a set function

In many presentations of the Lebesgue integral (although not in Lebesgue's original thesis) the integral is defined over arbitrary measurable sets E and denoted as

$$\int_E f(x) dx.$$

Then the integral over a compact interval $[a, b]$ would be written as

$$\int_{[a,b]} f(x) dx$$

and all of the theory is stated, as far as is possible, for the more general set-valued integral (rather than the interval-valued integral of this chapter). We can define this set-valued integral in somewhat greater generality by using estimates arising from full and fine covers.

Definition 7.37 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function and β a covering relation. We write

$$V(f\lambda, \beta) = \sup_{\pi \subset \beta} \left\{ \sum_{([u,v], w) \in \pi} |f(w)| \lambda([u, v]) \right\}$$

where the supremum is taken over all π , arbitrary subpartitions contained in β .

Definition 7.38 (Full and Fine Variations) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and let E be any set of real numbers. Then we define the full and fine variational measures associated with $f\lambda$ by the expressions:

$$V^*(f\lambda E) = \inf\{V(f\lambda, \beta) : \beta \text{ a full cover of } E\}$$

and

$$V_*(f\lambda, E) = \inf\{V(f\lambda, \beta) : \beta \text{ a fine cover of } E\}.$$

In the special case where f is a nonnegative function and E an arbitrary set we write

$$\int_E f(x) dx = V^*(f\lambda, E)$$

and we will check later to see if fine variation can be used as well. We have already sufficient techniques to study this set-valued integral and so we shall develop the theory in the exercises.

Exercises

Exercise 681 (measure estimates for Lebesgue's integral) Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ is an arbitrary nonnegative function and that $r < f(x) < s$ for all x in a set E . Then

$$r\lambda(E) \leq \int_E f(x) dx \leq s\lambda(E).$$

Answer $\square$

Exercise 682 (comparison with upper integral) Show that if f is a nonnegative function and E is an arbitrary set contained in an interval $[a, b]$ then

$$\int_E f(x) dx = \int_a^b \chi_E(x) f(x) dx.$$

$\square$

Exercise 683 (comparison with Lebesgue integral) Show that if f is a nonnegative measurable function and E is a measurable set contained in an interval $[a, b]$ then

$$\int_E f(x) dx = \int_a^b \chi_E(x) f(x) dx$$

7.9. THE LEBESGUE INTEGRAL AS A SET FUNCTION

301

where the integral may be interpreted as a Lebesgue integral. (In particular the value of the integral $\int_E f(x) dx$ can be constructed by Lebesgue's methods.) □

Exercise 684 (measure properties) Show that if f is a nonnegative function and $E, E_1, E_2, E_3, \dots$ is a sequence of sets for which $E \subset \bigcup_{n=1}^{\infty} E_n$ then

$$\int_E f(x) dx \leq \sum_{n=1}^{\infty} \int_{E_n} f(x) dx$$

i.e., the set function integral is a measure in the sense of Theorem 7.3. □

Exercise 685 (absolute continuity (zero/zero)) Show that if f is a nonnegative function and E is a set of Lebesgue measure zero then

$$\int_E f(x) dx = 0.$$

Answer □

Exercise 686 Show that if f is a nonnegative function and

$$\int_E f(x) dx = 0$$

then $f(x) = 0$ for almost every point $x \in E$. □

Exercise 687 Show that if f is a nonnegative function and $E_1, E_2, E_3, \dots$ is a sequence of pairwise disjoint closed sets for which $E = \bigcup_{n=1}^{\infty} E_n$ then

$$\int_E f(x) dx = \sum_{n=1}^{\infty} \int_{E_n} f(x) dx$$

i.e., the set function integral is additive over disjoint closed sets as in Corollary 7.9. □

Exercise 688 Suppose that f is a nonnegative, bounded function and that E is a measurable set. Show that for every $\varepsilon > 0$ there is an open set G so that $E \setminus G$ is closed and

$$\int_{E \setminus G} f(x) dx < \varepsilon.$$

[This is a warm-up to the next exercise where bounded is dropped.] □

Exercise 689 Suppose that f is a nonnegative, measurable function and that E is a measurable set. Show that for every $\varepsilon > 0$ there is an open set G so that $E \setminus G$ is closed and

$$\int_{E \setminus G} f(x) dx < \varepsilon.$$

[This is an improvement on the preceding exercise where it was assumed that the function is bounded.] □

Exercise 690 Show that if f is a nonnegative measurable function and $E_1, E_2, E_3, \dots$ is a sequence of pairwise disjoint measurable sets for which $E = \bigcup_{n=1}^{\infty} E_n$ then, for any set A ,

$$\int_{A \cap E} f(x) dx = \sum_{n=1}^{\infty} \int_{A \cap E_n} f(x) dx$$

i.e., the set function integral is additive over disjoint sets as in Lemma 7.12 provided we assume that the sets and the function are measurable. □

Exercise 691 Show that if f is a nonnegative measurable function and $E_1 \subset E_2 \subset E_3 \subset \dots$, is an increasing sequence of measurable sets for which $E = \bigcup_{n=1}^{\infty} E_n$ then

$$\int_E f(x) dx = \lim_{n \rightarrow \infty} \int_{E_n} f(x) dx.$$

□

Exercise 692 Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ and that f is nonnegative and bounded. Then for every $\varepsilon > 0$ there is a $\delta > 0$ so that if G is an open set with $\lambda(G) < \delta$ then

$$\int_G f(x) dx < \varepsilon.$$

[This is a warm-up to the next exercise where bounded is dropped.]

Answer □

Exercise 693 (absolute continuity (ε, δ)) Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$, that f is nonnegative and measurable, and that

$$\int_E f(x) dx < \infty.$$

Then for every $\varepsilon > 0$ there is a $\delta > 0$ so that if G is an open set with $\lambda(G) < \delta$ then

$$\int_{E \cap G} f(x) dx < \varepsilon.$$

Answer ☐

Exercise 694 (construction of the Lebesgue integral) Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ and that f is a nonnegative, measurable function. Let $r > 1$ and write

$$A_{kr} = \{x : r^{k-1} < f(x) \leq r^k\}.$$

Then, for any set E ,

$$\int_E f(x) dx \leq \sum_{k=-\infty}^{\infty} r^k \lambda(E \cap A_{kr}) \leq r \int_E f(x) dx.$$

[In particular as $r \nearrow 1$ the sum approaches the value of the integral.]

Answer ☐

Exercise 695 (full and fine characterization) Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ and that f is a nonnegative, measurable function. Show that

$$\int_E f(x) dx = V^*(f\lambda, E) = V_*(f\lambda, E).$$

Answer ☐

7.10 Characterizations of the indefinite integral

Under what conditions can we be sure that a function $F : [a, b] \rightarrow \mathbb{R}$ can be written as

$$F(t) = C + \int_a^t f(t) dt$$

for a constant C and an integrable function f . The property and the characterization itself for absolutely integrable functions were given by Giuseppe Vitali in 1905, only shortly after the publication by Lebesgue of his integration theory.

Definition 7.39 Suppose that $F : [a, b] \rightarrow \mathbb{R}$ is a function. Then F is absolutely continuous in the Vitali sense^a on $[a, b]$ if for all $\varepsilon > 0$ there is a $\delta > 0$ so that

$$\sum_i |F(v_i) - F(u_i)| < \varepsilon$$

whenever $\{[u_i, v_i]\}$ are nonoverlapping subintervals of $[a, b]$ for which $\sum_i [v_i - u_i] < \delta$.

^aMost texts call this (as did Vitali himself) “absolute continuity.” We prefer to reserve this term for the “zero variation on zero measure sets” which is the preferred use of the expression in measure theory.

There are several simple consequences of this definition that we will require in order to better understand this concept.

Lemma 7.40 Suppose that $F : [a, b] \rightarrow \mathbb{R}$ is a function that is absolutely continuous in the Vitali sense on $[a, b]$. Then

1. F is uniformly continuous on $[a, b]$,
2. F is absolutely continuous on (a, b) , and
3. F has bounded variation on $[a, b]$.

Proof. The first two statements are trivial and follow easily from the definition. For the third, choose a positive number δ so that

$$\sum_i |F(v_i) - F(u_i)| < 1$$

whenever $\{[u_i, v_i]\}$ are nonoverlapping subintervals of $[a, b]$ for which

$$\sum_i [v_i - u_i] < \delta.$$

Then any partition of $[a, b]$ into subintervals smaller than δ must have

$$\sum_i |F(v_i) - F(u_i)| < N$$

where N is an integer chosen large enough so that $N\delta > b - a$. ■

7.10.1 Integral of nonnegative, integrable functions

Theorem 7.41 Let $F : [a, b] \rightarrow \mathbb{R}$. A necessary and sufficient condition in order that F can be written as

$$F(t) = C + \int_a^t f(t) dt$$

for a constant C and a nonnegative integrable function f is that F is absolutely continuous in the Vitali sense and monotonic nondecreasing.

7.10.2 Integral of absolutely integrable functions

Theorem 7.42 Let $F : [a, b] \rightarrow \mathbb{R}$. A necessary and sufficient condition in order that F can be written as

$$F(t) = C + \int_a^t f(t) dt$$

for a constant C and an absolutely integrable function f is that F is absolutely continuous in the Vitali sense.

Corollary 7.43 Let $F : [a, b] \rightarrow \mathbb{R}$. A necessary and sufficient condition in order that F can be written as

$$F(t) = C + \int_a^t f(t) dt$$

for a constant C and an absolutely integrable function f is that

1. F is continuous on $[a, b]$.
2. F is absolutely continuous on (a, b) .
3. $V(F, [a, b]) < \infty$.

7.10.3 Integral of nonabsolutely integrable functions

Theorem 7.44 Let $F : [a, b] \rightarrow \mathbb{R}$. A necessary and sufficient condition in order that F can be written as

$$F(t) = C + \int_a^t f(t) dt$$

for a constant C and a nonabsolutely integrable function f are that

1. F is continuous on $[a, b]$.
2. F is absolutely continuous on (a, b) .
3. $V(F, [a, b]) = \infty$.
4. F is differentiable^a almost everywhere in (a, b) .

^aIt is possible but not easy to show that when F is absolutely continuous on (a, b) , F must be almost everywhere differentiable. Thus (4) follows from (3).

7.10.4 Proofs

The necessity of the conditions in the three theorems can be addressed first. Suppose that

$$F(t) = C + \int_a^t f(t) dt$$

for a constant C and an integrable function f .

If f is nonnegative then F is certainly nondecreasing. We check that it is also absolutely continuous in the Vitali sense.

Let $f_n(x) = \min\{f(x), n\}$ and note that f_n is measurable and nonnegative, and that $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ everywhere. Then, by the monotone convergence theorem, on every subinterval $[c, d] \subset [a, b]$,

$$0 < \int_c^d f(x) dx - \int_c^d f_n(x) dx < \int_c^d [f(x) - f_n(x)] dx \rightarrow 0.$$

Choose N so large that

$$\int_a^b f(x) dx < \int_a^b f_N(x) dx + \varepsilon/2.$$

Choose $\delta = \varepsilon/(2N)$. Then check that, if $[c_i, d_i]$ are nonoverlapping subintervals of $[a, b]$ with $\sum_i (d_i - c_i) < \delta$, then

$$\begin{aligned} 0 &\leq \sum_i [F(d_i) - F(c_i)] = \sum_i \int_{c_i}^{d_i} f(x) dx \\ &\leq \sum_i \int_{c_i}^{d_i} f_N(x) dx + \varepsilon/2 \\ &\leq \sum_i N((d_i - c_i) + \varepsilon/2) < N\delta + \varepsilon/2 < \varepsilon. \end{aligned}$$

This verifies that F is absolutely continuous in the Vitali sense.

If we assume instead that f is absolutely integrable we can again obtain the fact that F is absolutely continuous in the Vitali sense merely by splitting f into its positive and negative parts.

Finally, if f is merely integrable, then we already know that the relation

$$F(t) = C + \int_a^t f(t) dt$$

requires that F is continuous everywhere, and that F is absolutely continuous. The fundamental theorem of the calculus requires $F'(x) = f(x)$ almost everywhere in $[a, b]$. Thus each of the necessity parts of the three theorems is proved.

Conversely the stated conditions in the theorems are sufficient to verify that

$$F(t) = C + \int_a^t f(t) dt$$

for some function f as stated and constant C . For the third theorem we already know this from the fundamental theorem of the calculus.

That same theorem shows that the proof of the first theorem is also complete provided we know that F is differentiable almost everywhere and that $F'(x) \geq 0$ almost everywhere. But we already know that nondecreasing functions are almost everywhere differentiable. Take $f(x) = F'(x)$ at points where the derivative exists and $f(x) = 0$ elsewhere and the first theorem is proved.

We complete the proof of the second theorem in the same way. The assumption that F is absolutely continuous in the

Vitali sense assures us that F is continuous and has bounded variation. So again F is almost everywhere differentiable and again the same argument supplies the representation.

Exercises

Exercise 696 Show that a function that is absolutely continuous in the Vitali sense on $[a, b]$ must be uniformly continuous there. □

Exercise 697 Give an example of a uniformly continuous on an interval $[a, b]$ that is not absolutely continuous in the Vitali sense there. □

Exercise 698 Show that a function that is Lipschitz on $[a, b]$ is also absolutely continuous in the Vitali sense on $[a, b]$. □

Exercise 699 Given an example of a function that is not Lipschitz on $[a, b]$ but is absolutely continuous in the Vitali sense on $[a, b]$. □

Exercise 700 Show that a function that is absolutely continuous in the Vitali sense on $[a, b]$ must have bounded variation on $[a, b]$. □

Exercise 701 Show that if a function is absolutely continuous in the Vitali sense on $[a, b]$ then both parts of the Jordan decomposition have the same property on $[a, b]$. □

Exercise 702 Show that any continuously differentiable function on an interval $[a, b]$ is absolutely continuous in the Vitali sense on $[a, b]$. □

Exercise 703 Show that a differentiable function on an interval $[a, b]$ need not be absolutely continuous in the Vitali sense on $[a, b]$ but that it must be absolutely continuous in the more general sense (zero variation on zero measure sets). □

Exercise 704 Show that a function may be absolutely continuous but not absolutely continuous in the Vitali sense.

Answer ☐

Exercise 705 Let $F : \mathbb{R} \rightarrow \mathbb{R}$ and suppose that F is absolutely continuous in the Vitali sense on every compact interval $[a, b]$. Show that F is absolutely continuous.

Answer ☐

Exercise 706 Suppose that $F, f : [a, b] \rightarrow \mathbb{R}$, that f is bounded and integrable and that

$$F(t) = \int_a^b f(x) dx \quad (a \leq t \leq b).$$

Show directly that F is absolutely continuous in the Vitali sense on $[a, b]$.

Answer ☐

Exercise 707 Suppose that $F : [a, b] \rightarrow \mathbb{R}$ is absolutely continuous in $[a, b]$. Show that F is also absolutely continuous on $[a, b]$ in the sense of Vitali if and only if F has finite total variation on $[a, b]$, i.e., $V(F, [a, b]) < \infty$.

☐

Exercise 708 (Fichtenholz) Suppose that $F : [a, b] \rightarrow \mathbb{R}$ satisfies the following condition: for every $\varepsilon > 0$ there is a $\delta > 0$ so that whenever $\{[c_i, d_i]\}$ is any sequence of subintervals of $[a, b]$ satisfying $\sum_i (d_i - c_i) < \delta$ then necessarily $\sum_i |F(d_i) - F(c_i)| < \varepsilon$. Show that this condition is strictly stronger than absolute continuity in the Vitali sense.

Answer ☐

Exercise 709 Show that every Lipschitz function satisfies the condition of the preceding exercise.

☐

Exercise 710 Show that a function that satisfies the condition of the preceding exercises must be a Lipschitz function.

☐

7.11 Denjoy's program

For nonabsolutely integrable functions the integral is not constructive by any of the methods of Lebesgue. If we know in advance that $F'(x) = f(x)$ everywhere, then certainly we can “construct” the value of the integral by using the formula

$$\int_a^b f(x) dx = F(b) - F(a).$$

But even if we are assured that f is a derivative of some function, but we are not provided that function itself, then there may be no constructive method of determining either the value of the integral or the antiderivative function itself.

This may surprise some calculus students since much of an elementary course is devoted to various methods of *finding* antiderivatives.

After Lebesgue's constructive integral was presented there still remained this problem. All *bounded* derivatives can be handled by his methods, but there exist unbounded derivatives that are nonabsolutely integrable. What procedure (outside of our formal integration theory) would handle these?

Starting with the class of absolutely integrable functions, Arnaud Denjoy discovered in 1912 that a series of extensions of this class could be constructed that would eventually encompass all derivatives and, indeed, all nonabsolutely integrable functions. The methods are beyond the scope of this text as they require not merely an ordinary sequence of extensions, but a transfinite sequence of extensions using infinite ordinal numbers. He called his process totalization. Added to Lebesgue's methods, totalization reveals exactly how constructive our integral is. His process completely catalogues the class of nonabsolutely integrable functions. In effect the integral that is discussed in this text could be (and has been) called *the Denjoy integral*.

7.12 The Riemann integral

We conclude this chapter with a brief discussion of the Riemann integral. Since this has been used as the teaching integral of choice for many generations (in spite of criticisms) it can hardly be avoided. The student will surely encounter numerous references to it in the literature.

Definition 7.45 (Riemann integral) Suppose that f is an integrable function defined at every point of a compact interval $[a, b]$. Then f is said to be Riemann integrable on $[a, b]$ if for every $\varepsilon > 0$ there is a uniformly full cover β of $[a, b]$ so that

$$\left| \int_a^b f(x) dx - \sum_{[u,v] \in \pi} f(w)(v-u) \right| < \varepsilon$$

whenever π is a partition of the interval $[a, b]$ chosen from β .

Exercises

Exercise 711 Show that if f is almost everywhere continuous then f must be measurable. Deduce that if $f : [a, b] \rightarrow \mathbb{R}$ is bounded and almost everywhere continuous then f is Lebesgue integrable on $[a, b]$. $\square$

Exercise 712 Show that if f is bounded and almost everywhere continuous then f must satisfy McShane's criterion. Deduce that if $f : [a, b] \rightarrow \mathbb{R}$ is bounded and almost everywhere continuous then f is Lebesgue integrable on $[a, b]$. Answer $\square$

Exercise 713 Let $f : [a, b] \rightarrow \mathbb{R}$ be a bounded function. Prove that assertion (1) implies assertion (2):

1. For every $\varepsilon > 0$ there is a partition π of $[a, b]$ for which

$$\sum_{(I,x) \in \pi} \omega f(I) \lambda(I) < \varepsilon.$$

2. f is continuous at almost every point of $[a, b]$. Answer $\square$

Exercise 714 Let $f : [a, b] \rightarrow \mathbb{R}$ be a bounded function. Prove that assertion (2) implies assertion (1):

1. For every $\varepsilon > 0$ there is a partition π of $[a, b]$ for which

$$\sum_{(I,x) \in \pi} \omega f(I) \lambda(I) < \varepsilon.$$

2. f is continuous at almost every point of $[a, b]$. Answer $\square$

Exercise 715 (Lebesgue's criterion) Suppose that f is a function defined at every point of a compact interval $[a, b]$. Then f is Riemann integrable on $[a, b]$ if and only if f is bounded and almost everywhere continuous on (a, b) . Answer $\square$

Exercise 716 (Riemann's integrability criterion) Let $f : [a, b] \rightarrow \mathbb{R}$ be a bounded function. Then f is Riemann integrable if and only if for every $\varepsilon > 0$ there is a partition π of $[a, b]$ for which

$$\sum_{(I,x) \in \pi} \omega f(I) \lambda(I) < \varepsilon.$$

□

Exercise 717 A careless student argues: If a bounded function f is almost everywhere continuous that means that there is a continuous function g that is almost everywhere equal to f . Obviously this gives a much easier proof of Exercise 711. Your comments?

□

Chapter 8

Stieltjes Integrals

Recall that the *total variation* of a function F on a compact interval is the supremum of sums of the form

$$V(F, [a, b]) = \sum_{([u, v], w) \in \pi} |F(v) - F(u)|$$

taken over all possible partitions π of $[a, b]$. This is a measure of the variability of the function F on this interval. Functions of bounded variation play a significant role in real analysis. The earliest application was to the study of arc length of curves (see Section 3.9.3, a subject we will review in this chapter as well).

Our main tool in the study of this important class of functions is a slight generalization of the integral, called the Stieltjes integral. Our definitions for this integral will now be of the Henstock-Kurzweil type. Ideas related to the calculus integral will certainly return.

8.1 Stieltjes integrals

The definition of the total variation $V(F, [a, b])$ and the definition of the Lebesgue-Stieltjes measure both contain what looks very much like one of our Riemann sums, but in place of the usual sum

$$\sum_{([u, v], w) \in \pi} f(w)(v - u)$$

we are here checking values of the sum

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)|.$$

This might suggest to us that integration methods would prove a useful tool in the study of functions of bounded variation.

Let us, accordingly, enlarge the scope of our integration theory by considering limits of Riemann sums that are more general than we have used so far. Let $f, G : [a, b] \rightarrow \mathbb{R}$ and by analogy with

$$\int_a^b f(x) dx \sim \sum_{([u,v],w) \in \pi} f(w)(v - u)$$

we introduce new integrals by making only the obvious changes suggested by the following slogans:

$$\begin{aligned} \int_a^b f(x) dG(x) &\sim \sum_{([u,v],w) \in \pi} f(w)(G(v) - G(u)) \\ \int_a^b f(x) |dG(x)| &\sim \sum_{([u,v],w) \in \pi} f(w)|G(v) - G(u)| \\ \int_a^b f(x) [dG(x)]^+ &\sim \sum_{([u,v],w) \in \pi} f(w)[G(v) - G(u)]^+ \\ \int_a^b f(x) [dG(x)]^- &\sim \sum_{([u,v],w) \in \pi} f(w)[G(v) - G(u)]^- \end{aligned}$$

as well as a few other variants we consider in later sections:

$$\int_a^b \sqrt{|dG(x)| dx} \sim \sum_{([u,v],w) \in \pi} \sqrt{|G(v) - G(u)|(v - u)}$$

and

$$\int_a^b \sqrt{[dG(x)]^2 + [dx]^2} \sim \sum_{([u,v],w) \in \pi} \sqrt{|G(v) - G(u)|^2 + (v - u)^2}.$$

We will refer to all of these as *Stieltjes integrals*, although it is only the first variant of these,

$$\int_a^b f(x) dG(x),$$

that the Dutch mathematician Thomas Stieltjes (1856–1894) himself used and the one that most people would mean by the terminology.

8.1.1 Definition of the Stieltjes integral

The slogans in the preceding section should be enough to lead the reader to the correct definition of the various Stieltjes integral. Even so, let us give precise definitions for the simplest case. This is just a copying exercise: take the usual definition and repeat it with the Riemann sums adjusted in the manner required.

Definition 8.1 For functions $G, f : [a, b] \rightarrow \mathbb{R}$ we define an upper integral by

$$\overline{\int_a^b} f(x) dG(x) = \inf_{\beta} \sup_{\pi \subset \beta, ([u,v], w) \in \pi} \sum f(w)(G(v) - G(u))$$

where the supremum is taken over all partitions π of $[a, b]$ contained in β , and the infimum over all full covers β .

Similarly we define a *lower integral*, as

$$\underline{\int_a^b} f(x) dG(x) = \sup_{\beta} \inf_{\pi \subset \beta, ([u,v], w) \in \pi} \sum f(w)(G(v) - G(u))$$

where, again, π is a partition of $[a, b]$ and β is a full cover.

If the upper and lower integrals are identical we say the integral is *determined* and we write the common value as

$$\int_a^b f(x) dG(x).$$

We are interested, mostly, in the case in which the integral is determined and finite.

Exercises

Exercise 718 Let $G : [a, b] \rightarrow \mathbb{R}$. Show that

$$\int_a^b dG(x) = G(b) - G(a).$$

□

Exercise 719 Let $G : \mathbb{R} \rightarrow \mathbb{R}$ defined so that $G(x) = 0$ for all $x \neq 0$ and $G(0) = 1$. Compute

$$\int_0^2 |dG(x)| \quad \text{and} \quad \int_0^2 |dG(x)|.$$

□

Exercise 720 Let $G : [0, 1] \rightarrow \mathbb{R}$ and let $f(x) = 0$ for all $x \neq 1/2$ with $f(1/2) = 1$. What are

$$\int_0^1 f(x) dG(x) \quad \text{and} \quad \int_0^1 f(x) dG(x)?$$

□

Exercise 721 Let $G, f : [0, 1] \rightarrow \mathbb{R}$ and let $G(x) = 0$ for all $x \leq 1/2$ and with $G(x) = 1$ for all $x > 1/2$. What are

$$\int_0^1 f(x) dG(x) \quad \text{and} \quad \int_0^1 f(x) dG(x)?$$

Answer □

Exercise 722 Let $G, f : [a, b] \rightarrow \mathbb{R}$ and let f be continuous and let G be a step function, i.e. there are points

$$a < \xi_1 < \xi_2 < \cdots < \xi_m < b$$

so that G is constant on each interval (ξ_{i-1}, ξ_i) . What are possible values for

$$\int_a^b f(x) dG(x) \quad \text{and} \quad \int_a^b f(x) dG(x)?$$

Answer □

Exercise 723 Let $G, F : [-1, 1] \rightarrow \mathbb{R}$ be defined by $F(x) = 0$ for $-1 \leq x < 0$, $F(x) = 1$ for $0 \leq x \leq 1$, $G(x) = 0$ for $-1 \leq x \leq 0$, and $G(x) = 1$ for $0 < x \leq 1$. Discuss $\int_{-1}^1 F(x) dG(x)$ and $\int_{-1}^1 G(x) dF(x)$.

Answer □

Exercise 724 If $a < b < c$ is the formula

$$\int_a^b f(x) dG(x) + \int_b^c f(x) dG(x) = \int_a^c f(x) dG(x)$$

valid?

Answer ☐

Exercise 725 Show that a function f can be altered at a finite number of points where G is continuous without altering the values of the upper and lower integrals. Give an example to show that continuity may not be dropped here. ☐

Exercise 726 Show that a function f can be altered at a countable number of points where G is continuous without altering the values of the upper and lower integrals. ☐

Exercise 727 Give a Cauchy I criterion for $\int_a^b f(x) dG(x)$. ☐

Exercise 728 Give a Cauchy II criterion for $\int_a^b f(x) dG(x)$. ☐

Exercise 729 Give a McShane criterion for $\int_a^b f(x) dG(x)$. ☐

Exercise 730 Give a Henstock criterion for $\int_a^b f(x) dG(x)$. ☐

Exercise 731 For integrals of the form $\int_a^b f(x) |dG(x)|$ what changes have to be made in the various criteria?
 Answer ☐

Exercise 732 For integrals of the form $\int_a^b f(x) [dG(x)]^+$ what changes have to be made in the various criteria? ☐

Exercise 733 Let $F : [0, 2] \rightarrow \mathbb{R}$ with $F(t) = 0$ for all $t \neq 1$ and $F(1) = 1$. Show that

$$\int_0^2 |dF(x)| < \overline{\int_0^2 |dF(x)|} = V(F, [0, 2]).$$

☐

Exercise 734 Let $F : [a, b] \rightarrow \mathbb{R}$. Show that the total variation of F can be expressed as an upper integral:

$$V(F, [a, b]) = \overline{\int_a^b} |dF(x)|.$$

□

Exercise 735 Let $F : [a, b] \rightarrow \mathbb{R}$ and suppose that one at least of the integrals

$$\overline{\int_a^b} |dF(x)|, \quad \overline{\int_a^b} [dF(x)]^+ \quad \text{or} \quad \overline{\int_a^b} [dF(x)]^-$$

is finite. Show that F is a function of bounded variation on $[a, b]$ and that, for all $a < t \leq b$,

$$F(t) - F(a) = \overline{\int_a^t} [dF(x)]^+ - \overline{\int_a^t} [dF(x)]^-. \quad (8.1)$$

The identity (8.1) is a representation of F as a difference of two nondecreasing functions. □

Exercise 736 Let $F : [a, b] \rightarrow \mathbb{R}$ be a continuous function. Show that F has bounded variation on $[a, b]$ if and only if there is a continuous, strictly increasing function $G : [a, b] \rightarrow \mathbb{R}$ for which $F(d) - F(c) < G(d) - G(c)$ for all $a \leq c < d \leq b$. □

Exercise 737 What basic properties of the ordinary integral $\int_a^b f(x) dx$ from Chapter 6 can you prove for Stieltjes integrals without any but the most obvious of changes in the proofs? □

8.1.2 Henstock's zero variation criterion

Since the Stieltjes integral is defined in exactly the same way as the ordinary integral one expects almost the same properties. Indeed this integral has the same linear, additive, and monotone properties (suitably expressed). There also must be an indefinite integral. Finally, the most important of these properties that carries over, is the Henstock criterion. We give that now.

Theorem 8.2 Let $F, G, f : [a, b] \rightarrow \mathbb{R}$. Then a necessary and sufficient condition for the existence of the Stieltjes integral and the formula

$$\int_c^d f(x) dG(x) = F(d) - F(c) \quad [c, d] \subset [a, b]$$

is that

$$\int_a^b |dF(x) - f(x) dG(x)| = 0.$$

The proof would merely be a copying exercise of material from Chapter 6. Note that we are taking advantage of our general Stieltjes notation here to allow us to interpret the integral

$$\int_a^b |dF(x) - f(x) dG(x)|$$

as a limit of the Riemann sums

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u) - f(x)[G(v) - G(u)]|.$$

8.2 Regulated functions

Recall that the one-sided limit $F(c+)$ exists if, for all sequences of positive numbers t_n tending to zero,

$$\lim_{n \rightarrow \infty} F(c + t_n) = F(c+).$$

Similarly, we say $F(c-)$ exists if, for all sequences of positive numbers t_n tending to zero,

$$\lim_{n \rightarrow \infty} F(c - t_n) = F(c-).$$

Definition 8.3 Let $F : [a, b] \rightarrow \mathbb{R}$. Then

- F is said to be regulated if the one-sided limit $F(c+)$ exists and is finite for all $a \leq c < b$ and the limit on the other side $F(c-)$ exists and is finite for all $a < c \leq b$.
- F is said to be naturally regulated if F is regulated and, for all $a < c < b$, either

$$F(c+) \leq F(c) \leq F(c-)$$

or else

$$F(c-) \leq F(c) \leq F(c+).$$

Theorem 8.4 Let $F : [a, b] \rightarrow \mathbb{R}$ be monotonic. Then F is naturally regulated.

Proof. Simply notice that

$$\begin{aligned} F(c-) &= \sup\{F(t) : a \leq t < c\} \leq F(c) \\ &\leq \inf\{F(t) : c < t \leq b\} = F(c+). \end{aligned}$$

for all $a < c < b$. ■

Theorem 8.5 Let $F : [a, b] \rightarrow \mathbb{R}$ be a function of bounded variation. Then F is regulated and has at most countably many discontinuities^a.

^aIn fact it can be proved that all regulated functions have at most countably many discontinuities.

Proof. Suppose that $a < c \leq b$ and $F(c-)$ does not exist. Then there is a positive number ε and a sequence of numbers c_n increasing to c so that, for all n ,

$$F(c_n) - F(c_{n+1}) < -\varepsilon < \varepsilon < F(c_{n+2}) - F(c_{n+1}).$$

But then, for all m ,

$$\infty > V(F, [a, b]) \geq \sum_{n=1}^m |F(c_n) - F(c_{n+1})| > m\varepsilon.$$

This is impossible. Similarly $F(c+)$ must exist for all $a \leq c < b$.

Let us show that there are only countably many points $c \in [a, b]$ for which $F(c) \neq F(c+)$. Let $c_1, c_2, \dots, c_m$ denote some set of m points from (a, b) for which

$$|F(c_m+) - F(c)| > 1/n.$$

Then there is a disjointed collection of intervals $[c_i, t_i]$ for which

$$|F(t_i) - F(c_i)| > 1/(2n).$$

In particular

$$\infty > V(F, [a, b]) \geq \sum_{i=1}^m |F(t_i) - F(c_i)| > m/(2n).$$

Thus there are only finitely many such choices of points $c_1, c_2, \dots, c_m$ for which

$$|F(c_m+) - F(c_m)| > 1/n.$$

It follows that there are only countably many choices of points c_i for which

$$|F(c_i+) - F(c_i)| > 0.$$

A similar argument handles the points $c \in (a, b]$ for which $F(c) \neq F(c-)$. It follows that the set of points of discontinuity must be countable. ■

Lemma 8.6 (Approximate additivity) *Suppose that $F : [a, b] \rightarrow \mathbb{R}$ is a function that is naturally regulated. Then at any point $a < c < b$, and for any $\varepsilon > 0$ there is $\delta > 0$ so that, for all $c - \delta < u < c < v < c + \delta$,*

$$|F(v) - F(c)| + |F(c) - F(u)| \geq |F(v) - F(u)|$$

and

$$|F(v) - F(u)| \geq |F(v) - F(c)| + |F(c) - F(u)| - \varepsilon. \quad (8.2)$$

Proof. Since F is naturally regulated we know that

$$|F(c+) - F(c-)| = |F(c+) - F(c)| + |F(c-) - F(c)|$$

for each $a < c < b$. At such points there is a $\delta > 0$ so that

$$|F(u) - F(c-)| < \varepsilon/4 \text{ and } |F(v) - F(c+)| < \varepsilon/4$$

for all $c - \delta < u < c < v < c + \delta$. In particular

$$\begin{aligned} |F(c+) - F(c-)| &\leq |F(c+) - F(v)| + |F(v) - F(u)| + |F(u) - F(c-)| \\ &\leq |F(v) - F(u)| + \varepsilon/2 \end{aligned}$$

and so

$$\begin{aligned} |F(v) - F(c)| + |F(c) - F(u)| &\leq \\ |F(v) - F(c+)| + |F(c+) - F(c)| + |F(c-) - F(c)| + |F(c-) - F(u)| &\leq \\ \leq |F(c+) - F(c-)| + \varepsilon/2 \leq |F(v) - F(u)| + \varepsilon. \end{aligned}$$

Thus

$$|F(v) - F(u)| \geq |F(v) - F(c)| + |F(c) - F(u)| - \varepsilon.$$

The other inequality

$$|F(v) - F(c)| + |F(c) - F(u)| \geq |F(v) - F(u)|$$

is obviously true. ■

8.3 Variation expressed as an integral

We begin by pointing out the obvious relation between the Jordan variation and a certain Stieljtes integral.

Lemma 8.7 Suppose that $F : [a, b] \rightarrow \mathbb{R}$. Then

$$V(F, [a, b]) = \int_a^b |dF(x)|.$$

Our interest is in the special case where this integral exists and we are not forced to use the upper integral.

Lemma 8.8 Suppose that $F : [a, b] \rightarrow \mathbb{R}$ is a function of bounded variation that is naturally regulated. Then

$$V(F, [a, b]) = \int_a^b |dF(x)|.$$

Proof. It is clear that

$$V(F, [a, b]) \geq \int_a^b |dF(x)|.$$

In fact these are equal for all functions, but we do not need that. Let $\varepsilon > 0$ and select points

$$a = s_0 < s_1 < \cdots < s_{n-1} < s_n = b$$

so that

$$\sum_{i=1}^n |F(s_i) - F(s_{i-1})| > V(F, [a, b]) - \varepsilon.$$

Define a covering relation β to include only those pairs $([u, v], w)$ for which either $w \neq s_1, s_2, \dots, s_{n-1}$ and $[u, v]$ contains no point $s_1, s_2, \dots, s_{n-1}$, or else $w = s_i$ for some $i = 1, 2, \dots, n-1$ and

$$|F(v) - F(u)| \geq |F(v) - F(s_i)| + |F(s_i) - F(u)| - \varepsilon/n. \quad (8.3)$$

It is clear that β is full at every point w . For points $w \neq s_1, s_2, \dots, s_{n-1}$ this is transparent, while for points $w = s_i$ for some $i = 1, 2, \dots, n-1$, Lemma 8.6 may be applied.

We use a standard endpointed argument. Take any partition π of $[a, b]$ chosen from β . Scan through π looking for any elements of the form $([u, v], s_i)$ for $u < s_i < w$ and $i = 1, 2, \dots, n-1$. Replace each one by the new elements $([u, s_i], s_i)$ and $([s_i, v], s_i)$. Call the new partition π' . Because of (8.3) we see that

$$\sum_{([u, v], w) \in \pi} |F(v) - F(u)| \geq \sum_{([u, v], w) \in \pi'} |F(v) - F(u)| - \varepsilon.$$

Write $\pi_i = \pi'([s_{i-1}, s_i])$ and note that, by the way we have arranged π' , each π_i is a partition of the interval $[s_{i-1}, s_i]$. Consequently

$$\begin{aligned} \sum_{([u, v], w) \in \pi} |F(v) - F(u)| &\geq \sum_{([u, v], w) \in \pi'} |F(v) - F(u)| - \varepsilon \\ &\geq \sum_{i=1}^n \sum_{([u, v], w) \in \pi_i} |F(v) - F(u)| - \varepsilon \\ &\geq \sum_{i=1}^n |F(s_i) - F(s_{i-1})| - \varepsilon > V(F, [a, b]) - 2\varepsilon. \end{aligned}$$

We have shown that for every partition π of $[a, b]$ contained in β this sum is larger than $V(F, [a, b]) - 2\varepsilon$. It follows that

$$\int_a^b |dF(x)| \geq V(F, [a, b]) - 2\varepsilon.$$

Since ε is arbitrary the inequality

$$V(F, [a, b]) \leq \int_a^b |dF(x)| \leq \overline{\int_a^b |dF(x)|} \leq V(F, [a, b])$$

must hold and the theorem is proved. ■

Corollary 8.9 Suppose that $F : [a, b] \rightarrow \mathbb{R}$ is a function of bounded variation that is naturally regulated. Then

$$V(F, [a, b]) = \int_a^b |dF(x)| = \int_a^t [dF(x)]^+ + \int_a^t [dF(x)]^-.$$

Proof. The proof of the lemma can easily be adjusted to prove that all three of these integrals must exist. The identity is trivial: the expression

$$dF(x) = [dF(x)]^+ + [dF(x)]^-$$

integrated over $[a, b]$ produces the required identity. ■

The role of the naturally regulated assumption is exhibited in Exercise 733. It can be checked that if a function is not naturally regulated then the integral is not determined and the variation must be displayed using the upper integrals.

8.4 Representation theorems for functions of bounded variation

8.4.1 Jordan decomposition

The structure of functions of bounded variation is particularly simplified by a theorem of Jordan: every function of bounded variation is merely a linear combination of monotonic functions. We prove this for functions that are naturally regulated, by interpreting the statement as an integration assertion about certain Stieltjes integrals. The statement is true in general for all functions of bounded variation, but then the upper integrals would be needed (cf. Exercise 735).

Theorem 8.10 *Let $F : [a, b] \rightarrow \mathbb{R}$ be a function of bounded variation and suppose that F is naturally regulated. Then, for all $a < t \leq b$,*

$$F(t) - F(a) = \int_a^t [dF(x)]^+ - \int_a^t [dF(x)]^-. \quad (8.4)$$

The identity (8.4) is a representation of F as a difference of two functions, both nondecreasing, both naturally regulated.

Proof. The existence of the integrals is given in Corollary 8.9. The identity is trivial: the expression

$$dF(x) = [dF(x)]^+ - [dF(x)]^-$$

integrated over $[a, b]$ produces the required identity. ■

Corollary 8.11 *Let $F : [a, b] \rightarrow \mathbb{R}$ be a function of bounded variation and suppose that F is continuous. Then, for all $a < t \leq b$,*

$$F(t) - F(a) = \int_a^t [dF(x)]^+ - \int_a^t [dF(x)]^-. \quad (8.5)$$

The identity (8.5) is a representation of F as a difference of two functions, both continuous and nondecreasing.

8.4.2 Jordan decomposition theorem: differentiation

We know that all functions of bounded variation and all monotonic functions are almost everywhere differentiable. This and the integral representation given in Theorem 8.10 allows the following corollary.

Corollary 8.12 Let $F : [a, b] \rightarrow \mathbb{R}$ be a function of bounded variation and suppose that F is naturally regulated. Write

$$F_1(t) = \int_a^t [dF(x)]^+ \quad (a \leq t \leq b), \quad (8.6)$$

and

$$F_2(t) = \int_a^t [dF(x)]^- \quad (a \leq t \leq b), \quad (8.7)$$

Then

$$F(t) - F(a) = F_1(t) - F_2(t) \quad \text{and} \quad T(t) = V(F, [a, t]) = F_1(t) + F_2(t).$$

Moreover, at almost every t in $[a, b]$,

$$\begin{aligned} F'(t) &= F'_1(t) - F'_2(t), \quad F'_1(t) = \max\{F'(t), 0\}, \quad F'_2(t) = \max\{-F'(t), 0\}, \\ T'(t) &= F'_1(t) + F'_2(t) = |F'(t)| \quad \text{and} \quad F'_1(t)F'_2(t) = 0. \end{aligned}$$

Proof. There are three tools needed for the differentiation statements: the Lebesgue differentiation theorem (that monotonic functions have derivatives a.e.), the Henstock zero variation criterion for integrals, and the zero variation implies zero derivative a.e. rule.

We illustrate with a proof for one of the statements in the corollary. Define

$$h([u, v], w) = F_1(v) - F_1(u) - [F(v) - F(u)]^+.$$

The identity $F_1(t) = \int_a^t [dF(x)]^+$ requires that h have zero variation on (a, b) . This, in term, requires that

$$\begin{aligned} & \lim_{h \rightarrow 0+} \frac{F_1(t+h) - F_1(t) - \max\{F(t+h) - F(t), 0\}}{h} \\ &= \lim_{h \rightarrow 0+} \frac{F_1(t) - F_1(t-h) - \max\{F(t) - F(t-h), 0\}}{h} = 0 \end{aligned}$$

for almost every t in (a, b) . From that we deduce that $F'_1(t) = \max\{F'(t), 0\}$ must be true for almost every t in (a, b) . Proofs for the other statements are similar. ■

8.4.3 Representation by saltus functions

Theorem 8.13 Let $F : [a, b] \rightarrow \mathbb{R}$ be a monotonic nondecreasing function and let C be the set of points of continuity of F in $[a, b]$. Then, for all $a < t \leq b$,

$$F(t) - F(a) = \int_a^t \chi_C(x) dF(x) + \int_a^t [1 - \chi_C(x)] dF(x). \quad (8.8)$$

and

$$\int_a^t [1 - \chi_C(x)] dF(x) = [F(t) - F(t-)] + \sum_{s \in [a, t) \setminus C} [F(s+) - F(s-)]$$

The identity (8.8) is a representation of F as a sum of two functions, the first continuous and nondecreasing, the second a saltus function.

8.4.4 Representation by singular functions

Theorem 8.14 Let $F : [a, b] \rightarrow \mathbb{R}$ be a continuous monotonic function. Let D be the set of points of differentiability of F in $[a, b]$. Then

$$F(t) - F(a) = \int_a^t \chi_D(x) dF(x) + \int_a^t [1 - \chi_D(x)] dF(x) \quad (8.9)$$

and

$$\int_a^t \chi_D(x) dF(x) = \int_a^t F'(x) dx.$$

The identity (8.9) is a representation of F as a sum of two monotonic functions, the first Vitali absolutely continuous and the second a continuous singular function.

8.5 Reducing a Stieltjes integral to an ordinary integral

The Stieltjes integral reduces to an ordinary integral in a number of interpretations. When the integrating function G is an indefinite integral the whole theory reduces to ordinary integration. The formula is compelling since, as calculus students often learn,

$$dG(x) = G'(x) dx$$

can be assigned a meaning. That meaning is convenient here too and suggests that

$$\int_a^b f(x) dG(x) = \int_a^b f(x) G'(x) dx.$$

Theorem 8.15 Suppose that $G, f, g : \mathbb{R} \rightarrow \mathbb{R}$ and that g is integrable on a compact interval $[a, b]$ with an indefinite integral

$$G(d) - G(c) = \int_c^d g(x) dx \quad (a \leq c < d \leq b).$$

Then the Stieltjes integral

$$\int_a^b f(x) dG(x)$$

exists if and only if fg is integrable on $[a, b]$, in which case

$$\int_a^b f(x) dG(x) = \int_a^b f(x)g(x) dx.$$

Proof. The proof depends simply on the Henstock criterion. The existence of the ordinary integral

$$\int_a^b g(x) dx$$

with an indefinite integral G is equivalent to the zero criterion:

$$\int_a^b |dG(x) - g(x) dx| = 0$$

Whenever this identity holds, then one checks that, for any function f ,

$$\int_a^b |f(x)dG(x) - f(x)g(x) dx| = 0$$

would also be true. For example, if we have a bounded f this is trivial; for unbounded one only has to split $[a, b]$ into the sequence of sets

$$\{x \in [a, b] : n-1 \leq |f(x)| < n\}$$

and argue on each of these (cf. Exercise 739).

The existence of the Stieltjes integral

$$\int_a^b f(x) dG(x)$$

with an indefinite integral F is equivalent to the zero criterion:

$$\int_a^b |dF(x) - f(x) dG(x)| = 0.$$

Together these give

$$\begin{aligned} \int_a^b |dF(x) - f(x)g(x) dx| &\leq \\ \int_a^b |dF(x) - f(x) dG(x)| + \int_a^b |f(x)dG(x) - f(x)g(x) dx| &= 0. \end{aligned}$$

From this it is easy to read off the required identity. ■

8.6 Properties of the indefinite integral

Theorem 8.16 Suppose that

$$F(t) = \int_a^t f(x) dG(x) \quad (a \leq t \leq b).$$

Then

1. F is continuous at every point at which G is continuous.
2. F is absolutely continuous in any set $E \subset (a, b)$ in which G is absolutely continuous.
3. F has zero variation on any set $E \subset (a, b)$ on which G has zero variation.
4. F has bounded variation on $[a, b]$ if f is bounded and if G has bounded variation.
5. If G is Vitali absolutely continuous on $[a, b]$ and if f is bounded then F is also Vitali absolutely continuous on $[a, b]$.
6. If G is a saltus function on $[a, b]$ and f is nonnegative then so too is the indefinite integral F . Moreover the jumps of F occur precisely at points that are jumps of G for which f does not vanish.

Theorem 8.17 (Differentiation properties) Suppose that

$$F(t) = \int_a^t f(x) dG(x) \quad (a \leq t \leq b).$$

Then

1. For almost every point x in $[a, b]$

$$\lim_{y \rightarrow x} \frac{F(y) - F(x) - f(x)(G(y) - G(x))}{y - x} = 0.$$

2. For almost every point x in $[a, b]$,

$$\overline{DF}(x) = f(x)\overline{DG}(x) \quad \text{and} \quad \underline{DF}(x) = f(x)\underline{DG}(x)$$

or else

$$\overline{DF}(x) = f(x)\underline{DG}(x) \quad \text{and} \quad \underline{DF}(x) = f(x)\overline{DG}(x)$$

depending on whether $f(x) \geq 0$ or $f(x) \leq 0$.

3. In particular, $F'(x) = f(x)G'(x)$ at almost every point x at which either F or G is differentiable.
4. Finally, $F'(x) = 0$ at almost every point x where $f(x) = 0$.

The proof for each of these statements depends simply on the Henstock criterion. The existence of the Stieltjes integral

$$\int_a^b f(x) dG(x)$$

with an indefinite integral F is equivalent to the zero criterion:

$$\int_a^b |dF(x) - f(x) dG(x)| = 0$$

From the latter will flow each of the statements of the theorem. The individual proofs are left in the Exercises to the reader.

Exercises

Exercise 738 Suppose that

$$\int_a^b |dF(x) - f(x) dx| = 0.$$

Show that if g is any bounded function on $[a, b]$ then

$$\int_a^b |g(x)dF(x) - f(x)g(x) dx| = 0.$$

□

Exercise 739 Suppose that

$$\int_a^b |dF(x) - f(x) dx| = 0.$$

Show that if g is any real-valued function on $[a, b]$ then

$$\int_a^b |g(x)dF(x) - f(x)g(x) dx| = 0.$$

□

Exercise 740 Suppose that

$$\int_a^b |dF(x) - f(x) dG(x)| = 0.$$

Show that F is continuous at any point at which G is continuous. Is the converse necessarily true?

□

Exercise 741 Suppose that

$$\int_a^b |dF(x) - f(x) dG(x)| = 0.$$

Show that F has zero variation on any set on which G has zero variation. Is the converse necessarily true?

□

Exercise 742 Suppose that

$$\int_a^b |dF(x) - f(x) dG(x)| = 0$$

8.6. PROPERTIES OF THE INDEFINITE INTEGRAL

333

and suppose that G has bounded variation on $[a, b]$ and that f is bounded. Show that F has bounded variation on $[a, b]$. □

Exercise 743 Suppose that

$$\int_a^b |dF(x) - f(x) dG(x)| = 0.$$

Show that

$$\lim_{y \rightarrow x} \frac{F(y) - F(x) - f(x)(G(y) - G(x))}{y - x} = 0$$

almost everywhere by using the zero variation implies zero derivative criterion. □

Exercise 744 Complete the remaining arguments needed to establish the parts of the theorem. □

Exercise 745 Suppose that

$$\int_a^b |dF(x) - f(x) dG(x)| = 0.$$

Show that, for every point x in $[a, b]$

$$\lim_{y \rightarrow x} \frac{F(y) - F(x)}{G(y) - G(x)} = f(x)$$

except perhaps for points x in a set N in which G has fine variation zero. □

Exercise 746 Suppose that at every point x of a compact interval $[a, b]$

$$\lim_{y \rightarrow x} \frac{F(y) - F(x) - f(x)[G(y) - G(x)]}{y - x} = 0.$$

Show that

$$\int_a^b |dF(x) - f(x) dG(x)| = 0.$$

□

Exercise 747 Suppose that at every point x of a compact interval $[a, b]$

$$\lim_{y \rightarrow x} \frac{F(y) - F(x) - f(x)[G(y) - G(x)]}{y - x} = 0$$

except for points x in a set N for which both F and G have zero variation. Show that

$$\int_a^b |dF(x) - f(x) dG(x)| = 0.$$

□

Exercise 748 Suppose that

$$\int_a^b |dF(x) - f(x) dG(x)| = 0.$$

Show that, at almost every point x ,

$$\overline{DF}(x) = f(x)\overline{DG}(x) \text{ and } \underline{DF}(x) = f(x)\underline{DG}(x)$$

if $f(x) \geq 0$ while

$$\overline{DF}(x) = f(x)\underline{DG}(x) \text{ and } \underline{DF}(x) = f(x)\overline{DG}(x)$$

if $f(x) \leq 0$. In particular $F'(x) = 0$ at almost every point x where $f(x) = 0$.

□

8.6.1 Existence of the integral from derivative statements

The existence of the integral

$$\int_a^b f(x) dG(x)$$

can be deduced from a variety of differentiation statements. For example, using Exercise 747, we can prove the following simple version:

Theorem 8.18 Suppose that at every point x of a compact interval $[a, b]$

$$\lim_{y \rightarrow x} \frac{F(y) - F(x) - f(x)[G(y) - G(x)]}{y - x} = 0$$

except for points x in a set N for which both F and G have zero variation. Then the Stieltjes integral exists and

$$\int_a^b f(x) dG(x) = F(b) - F(a).$$

8.7 Existence of the Stieltjes integral for continuous functions

Theorem 8.19 Let $f, G : \mathbb{R} \rightarrow \mathbb{R}$ and suppose that f is continuous on a compact interval $[a, b]$ and that G is monotonic nondecreasing throughout that interval. Then the Stieltjes integral exists and

$$\left| \int_a^b f(x) dG(x) \right| \leq \|f\|_{\infty} [G(b) - G(a)].$$

where $\|f\|_{\infty} = \max_{t \in [a, b]} |f(t)|$.

Proof. The inequality is easy since, for any pair $([u, v], w)$ with $[u, v] \subset [a, b]$,

$$|f(w)(G(v) - G(u))| \leq \|f\|_{\infty} [G(v) - G(u)]. \quad (8.10)$$

To prove that the integral exists we can invoke a version of the McShane criterion here. The details are left as an exercise. ■

The next theorem is similar.

Theorem 8.20 Let $f, G : \mathbb{R} \rightarrow \mathbb{R}$ and suppose that f is continuous on a compact interval $[a, b]$ and that G has bounded variation throughout that interval. Then the Stieltjes integral exists and

$$\left| \int_a^b f(x) dG(x) \right| \leq \|f\|_{\infty} V(G, [a, b]).$$

where $\|f\|_{\infty} = \max_{t \in [a, b]} |f(t)|$.

8.8 Integration by parts

Integration by parts for the Stieltjes integral assumes the following form¹:

Theorem 8.21 *Let $F, G : \mathbb{R} \rightarrow \mathbb{R}$. Then*

$$\int_a^b [F(x) dG(x) + G(x) dF(x)] = F(b)G(b) - F(a)G(a) - \int_a^b dF(x) dG(x)$$

in the sense that if one of the integrals exists, so too does the other with the stated identity.

Proof. First check a simple identity: that, for any u and v ,

$$\begin{aligned} & F(u)[G(v) - G(u)] + G(u)[F(v) - F(u)] \\ &= F(v)G(v) - G(u)G(u) - [F(v) - F(u)][G(v) - G(u)]. \end{aligned}$$

This suggests that

$$\int_a^b |F(x) dG(x) + G(x) dF(x) - dF(x) dG(x) - dF(x) dG(x)| = 0 \quad (8.11)$$

is simply true because of an identity. If indeed this is true then the statement in the theorem is obvious because

$$\int_a^b dF(x) dG(x) = F(b)G(b) - F(a)G(a).$$

To complete the proof we have to address just one concern here. If a partition π of the interval $[a, b]$ contains only pairs $([u, v], u)$ or $([u, v], v)$ [i.e., $([u, w], w)$ with w only at an endpoint] then our simple identity would indeed supply

$$\begin{aligned} & \sum_{([u, v], w) \in \pi} [F(w)[G(v) - G(u)] + G(w)[F(v) - F(u)] - F(v)G(v) - G(u)G(u) \\ &= \sum_{([u, v], w) \in \pi} [F(v) - F(u)][G(v) - G(u)]. \end{aligned}$$

That surely proves (8.11) if we are allowed to use only such partitions. But what happens if we permit (as we must) partitions π containing a pair $([u, v], w) \in \pi$ for which $u < w < v$?

¹For the Riemann-Stieltjes integral the extra term $\int_a^b dF(x) dG(x)$ does not appear, since this would be zero whenever the integral exists in that sense. (See Corollary 8.23, which should look familiar to fans of the Riemann-Stieltjes integral.)

To clear this up note that we can always adjust full covers and partitions π by replacing any pair $([u, v], w) \in \pi$ for which $u < w < v$ by the two items $([u, w], w)$ and $([w, v], w)$. That does not change the sums here because, for example,

$$F(w)[G(v) - G(u)] = F(w)[G(w) - G(u)] + F(w)[G(v) - G(w)].$$

This “endpointed” argument (which we have seen before in Exercise 649) means that in these simple Stieltjes integrals the partitions used can all be restricted to ones where only elements of the form $([u, v], u)$ or $([u, v], v)$ can appear. ■

Corollary 8.22 *Let $F, G : \mathbb{R} \rightarrow \mathbb{R}$ and suppose that*

$$\int_a^b |dF(x) dG(x)| = 0.$$

Then

$$\int_a^b [F(x) dG(x) + G(x) dF(x)] = F(b)G(b) - F(a)G(a).$$

If, in addition one of the following two integrals exists then so too does the other and

$$\int_a^b F(x) dG(x) + \int_a^b G(x) dF(x) = F(b)G(b) - F(a)G(a).$$

Corollary 8.23 *Let $F, G : \mathbb{R} \rightarrow \mathbb{R}$ and suppose that F is continuous and G has bounded variation. Then*

$$\int_a^b F(x) dG(x) + \int_a^b G(x) dF(x) = F(b)G(b) - F(a)G(a).$$

Proof. The assumption that F is continuous and G has bounded variation requires that

$$\int_a^b |dF(x) dG(x)| = 0.$$

Thus Theorem 8.21 can be applied. But we know, from Theorem 8.19, that the integral $\int_a^b F(x) dG(x)$ must exist. It follows, from Corollary 8.22, that $\int_a^b G(x) dF(x)$ must also exist and that the integration by parts formula is valid. ■

8.9 Lebesgue-Stieltjes measure

The variation of a function F on an interval $[a, b]$ can be described by the identity

$$V(F, [a, b]) = \sup_{([u, v], w) \in \pi} \sum |F(v) - F(u)|$$

where the supremum is taken over all possible partitions π of the interval $[a, b]$. We recall that a somewhat similar expression describes the Lebesgue measure $\lambda(E)$ of a set E :

$$\lambda(E) = \inf_{\beta} \sup_{\pi \in \beta} \sum_{([u, v], w) \in \pi} (v - u).$$

Here π denotes an arbitrary subpartition contained in β and the infimum is taken over all full covers β of the set E . There is an obvious generalization of Lebesgue measure available by replacing $(v - u)$ by $|F(v) - F(u)|$.

Definition 8.24 Let F be a function defined at least on an open set G and we suppose that $E \subset G$. Then we write

$$\lambda_F(E) = \inf_{\beta} \sup_{\pi \in \beta} \sum_{([u, v], w) \in \pi} |F(v) - F(u)|.$$

Here π denotes an arbitrary subpartition contained in β . The set function λ_F defined for all subsets of G is called the Lebesgue-Stieltjes measure associated with F or, often, the variational measure associated with F .

In the literature often the Lebesgue-Stieltjes measure is studied only for monotonic functions that are continuous on the left-hand side at every point. It is convenient for us to usurp this language for the completely general case. The definition of the Lebesgue-Stieltjes measure is closely related to the Stieltjes integral, just as the definition of Lebesgue measure in Lemma 7.2 was expressible as an upper integral.

Lemma 8.25 If F is defined on a compact interval $[a, b]$ and $E \subset (a, b)$ then

$$\lambda_F(E) = \int_a^{\overline{b}} \chi_E(x) |dF(x)|.$$

By comparing this definition with some earlier notions that are almost identical we will be able to deduce the following properties of this measure:

Properties of the Lebesgue-Stieltjes measures

1. λ_F is a measure, i.e., if F is defined on an open set G and $E, E_1, E_2, E_3, \dots$ are subsets of G for which $E \subset \bigcup_{n=1}^{\infty} E_n$ then this inequality must hold:

$$\lambda_F(E) \leq \sum_{n=1}^{\infty} \lambda_F(E_n).$$

2. If F is monotonic then

$$\lambda_F([a, b]) = |F(b+) - F(a-)|,$$

$$\lambda_F((a, b)) = |F(b-) - F(a+)|,$$

and

$$\lambda_F(\{x_0\}) = |F(x_0+) - F(x_0-)|.$$

3. F has zero variation on a set E if and only if $\lambda_F(E) = 0$.
4. F is continuous at a point x_0 if and only if $\lambda_F(\{x_0\}) = 0$.
5. F is continuous at every point of an open interval (a, b) if and only if $\lambda_F(C) = 0$ for every countable subset of (a, b) .
6. F is absolutely continuous on an interval (a, b) if and only if $\lambda_F(N) = 0$ for every subset N of (a, b) that has measure zero.
7. $\lambda_F((a, b)) = 0$ if and only if F is constant on (a, b) .
8. F is locally bounded at a point x_0 if and only if $\lambda_F(\{x_0\}) < \infty$.
9. If F is defined on a compact interval $[a, b]$ then F has bounded variation on $[a, b]$ if and only if $\lambda_F((a, b)) < \infty$.
10. If F is defined on an open set G and has a bounded derivative at each point of a bounded subset E of G then $\lambda_F(E) < \infty$.
11. If F is defined on an open set G and $\lambda_F(E) < \infty$ then F is differentiable at almost every point of E .

It is clear from the definitions that F has zero variation on a set E if and only if $\lambda_F(E) = 0$. Thus the assertions (4)–(8) are immediate from our early study of zero variation. The other assertions are proved in the exercises.

Exercises

Exercise 749 Prove that λ_F is a measure.

Answer ☐

Exercise 750 Show that if F is monotonic then F is monotonic then

$$\lambda_F([a, b]) = |F(b+) - F(a-)|,$$

$$\lambda_F((a, b)) = |F(b-) - F(a+)|,$$

and

$$\lambda_F(\{x_0\}) = |F(x_0+) - F(x_0-)|.$$

☐

Exercise 751 Show that, if the one-sided limits $F(x_0+)$ and $F(x_0-)$ exist then

$$\lambda_F(\{x_0\}) = |F(x_0+) - F(x_0)| + |F(x_0-) - F(x_0)|.$$

☐

Exercise 752 Suppose that F is defined on an open set G . Show that F is locally bounded at a point $x_0 \in G$ if and only if $\lambda_F(\{x_0\}) < \infty$.

☐

Exercise 753 Suppose that F is defined on a compact interval $[a, b]$. Show that F has bounded variation on $[a, b]$ if and only if $\lambda_F((a, b)) < \infty$. Show that $\lambda_F((a, b)) \leq V(F, [a, b])$ but that the inequality may be strict unless F is continuous.

Answer ☐

Exercise 754 Suppose that F is defined on an open set G and has a bounded derivative at each point of a bounded subset E of G . Show that $\lambda_F(E) < \infty$.

Answer ☐

5We recall that every function of bounded variation is

8.10 Mutually singular functions

Definition 8.26 Let $F, G : [a, b] \rightarrow \mathbb{R}$ be functions of bounded variation. Then F and G are said to be mutually singular provided that

$$\int_a^b \sqrt{|dF(x) dG(x)|} = 0.$$

Lemma 8.27 Let $F, G : [a, b] \rightarrow \mathbb{R}$ be functions of bounded variation. If F and G are mutually singular, then $F'(x)G'(x) = 0$ almost everywhere in $[a, b]$.

Proof. This follows easily (as usual) from the zero variation implies zero derivative a.e. rule together with the fact that both $F'(x)$ and $G'(x)$ must exist a.e.. ■

Our main theorem shows that mutually singular functions grow on separate parts of the interval $[a, b]$ in a sense made precise here.

Theorem 8.28 Let $F, G : [a, b] \rightarrow \mathbb{R}$ be functions of bounded variation. Then F and G are mutually singular on $[a, b]$ if and only for every $\varepsilon > 0$ there is a full cover β with the property that every partition π of $[a, b]$ contained in β can be split into two disjoint subpartitions $\pi = \pi' \cup \pi''$ so that

$$\sum_{([u,v],w) \in \pi'} |F(v) - F(u)| < \varepsilon$$

and

$$\sum_{([u,v],w) \in \pi''} |G(v) - G(u)| < \varepsilon.$$

Proof. Suppose that

$$\int_a^b \sqrt{|dF(x) dG(x)|} = 0.$$

Let $\varepsilon > 0$ and select a full cover β so that

$$\sum_{([u,v],w) \in \pi} \sqrt{|[F(v) - F(u)][G(v) - G(u)]|} < \varepsilon$$

for all partitions π of $[a, b]$ contained in β . Split such a π as follows:

$$\pi' = \{([u, v], w) : |[F(v) - F(u)]| \leq |[G(v) - G(u)]|\}$$

and

$$\pi'' = \{([u, v], w) : |[F(v) - F(u)]| > |[G(v) - G(u)]|\}.$$

Verify that $\pi = \pi' \cup \pi''$ and that

$$\sum_{([u, v], w) \in \pi'} |[F(v) - F(u)]| \leq \sum_{([u, v], w) \in \pi'} \sqrt{|[F(v) - F(u)][G(v) - G(u)]|} < \varepsilon$$

and that

$$\sum_{([u, v], w) \in \pi''} |[G(v) - G(u)]| \leq \sum_{([u, v], w) \in \pi''} \sqrt{|[F(v) - F(u)][G(v) - G(u)]|} < \varepsilon.$$

This proves one direction in the theorem.

For the converse select a number $M > 0$ and a full cover β_1 so that

$$\sum_{([u, v], w) \in \pi} |[F(v) - F(u)]| + |[G(v) - G(u)]| < M$$

for all partitions π of $[a, b]$ from β_1 . This is possible merely because the functions F and G have bounded variation. Select a full cover β_2 with the property presented in the statement of the theorem (for ε). Let $\beta = \beta_1 \cap \beta_2$. This is a full cover. Consider any partition π of $[a, b]$ contained in β . There must be, by hypothesis, a split $\pi = \pi' \cup \pi''$ so that

$$\sum_{([u, v], w) \in \pi'} |[F(v) - F(u)]| < \varepsilon$$

and

$$\sum_{([u, v], w) \in \pi''} |[G(v) - G(u)]| < \varepsilon.$$

We now compute

$$\begin{aligned} \sum_{([u, v], w) \in \pi} \sqrt{|[F(v) - F(u)][G(v) - G(u)]|} &= \\ \sum_{([u, v], w) \in \pi'} \sqrt{|[F(v) - F(u)][G(v) - G(u)]|} &= \end{aligned}$$

$$\begin{aligned}
 & + \sum_{([u,v],w) \in \pi''} \sqrt{|[F(v) - F(u)][G(v) - G(u)]|} \\
 & \leq \sqrt{\sum_{([u,v],w) \in \pi'} |[F(v) - F(u)]|} \sqrt{\sum_{([u,v],w) \in \pi'} |[G(v) - G(u)]|} \\
 & + \sqrt{\sum_{([u,v],w) \in \pi''} |[F(v) - F(u)]|} \sqrt{\sum_{([u,v],w) \in \pi''} |[G(v) - G(u)]|} \\
 & \leq 2\sqrt{M\varepsilon}.
 \end{aligned}$$

Here we have used the Cauchy-Schwartz inequality. Since ε is an arbitrary positive number it follows that

$$\int_a^b \sqrt{|dF(x) dG(x)|} = 0.$$

Consequently F and G must be mutually singular. ■

8.11 Singular functions

We have defined the notion of a singular function elsewhere and given the usual remarkable example of such a function, the Cantor function (Devil's staircase). We show that there are further characterizations of this notion, in particular one given exactly by a Stieltjes-type integral.

Theorem 8.29 *Let $F : [a, b] \rightarrow \mathbb{R}$ be a function of bounded variation. Then the following are equivalent:*

1. F is singular.
2. $F'(x) = 0$ almost everywhere in $[a, b]$.
3. $\int_a^b \sqrt{|dF(x)|} dx = 0$.

Proof. It is only the third property that we show here, since we know from elsewhere that the first two are equivalent. If the third statement is true then we can check, using the zero variation implies zero derivative a.e. rule that $F'(x) = 0$ a.e..

Conversely suppose that $F'(x) = 0$ almost everywhere. Let $\varepsilon > 0$ and choose a sequence of open intervals $\{(c_i, d_i)\}$ with total length smaller than ε so that $F'(x) = 0$ for all $x \in [a, b]$ not in one of the intervals. Define two covering relations. The first β_1 consists of all pairs $([u, v], w)$ subject only to the condition that if w is in $[a, b]$ and not covered by an open interval $\{(c_i, d_i)\}$ then

$$|F(v) - F(u)| < \varepsilon(v - u)/(b - a).$$

The second β_2 consists of all pairs $([u, v], w)$ subject only to the condition that if w is contained in one of the open intervals $\{(c_i, d_i)\}$ then so too is $[u, v]$. Then β_1 , β_2 , and $\beta = \beta_1 \cap \beta_2$ are all full covers.

Note that if π is a subpartition contained in β_1 consisting of pairs $([u, v], w)$ not covered by an open interval from $\{(c_i, d_i)\}$ then

$$\sum_{([u, v], w) \in \pi} |F(v) - F(u)| \leq \sum_{([u, v], w) \in \pi} \varepsilon(v - u)/(b - a) \leq \varepsilon.$$

Note that if π is a subpartition contained in β_2 consisting of pairs $([u, v], w)$ that are covered by an open interval from $\{(c_i, d_i)\}$ then

$$\sum_{(I, x) \in \pi} (v - u) \leq \sum_{i=1}^{\infty} (d_i - c_i) < \varepsilon.$$

Thus any partition of $[a, b]$ chosen from β can be split into two subpartitions with these inequalities. This verifies the conditions asserted in Theorem 8.28 for F and the function $G(x) = x$. But that is exactly our third condition in the statement of the theorem. ■

8.12 Length of curves

A curve is a pair of continuous functions $F, G : [a, b] \rightarrow \mathbb{R}$. We consider that the curve is the pair of functions itself, rather than that the curve is the geometric set of points

$$\{(F(t), G(t)) : t \in [a, b]\}$$

that is the object we might likely think about when contemplating a curve.

Definition 8.30 Suppose that $F, G : [a, b] \rightarrow \mathbb{R}$ is a pair of continuous functions. By the length of the curve given by the pair F and G we shall mean

$$\int_a^b \sqrt{[dF(x)]^2 + [dG(x)]^2}.$$

That this integral is determined (but may be infinite) is pointed out in the proof of the next theorem.

Theorem 8.31 A curve given by a pair of continuous functions $F, G : [a, b] \rightarrow \mathbb{R}$ has finite length if and only if both functions F and G have bounded variation.

Proof. Note that as F and G are continuous, then so too is the interval function

$$h([u, v]) = \sqrt{[F(v) - F(u)]^2 + [G(v) - G(u)]^2}.$$

A simple application of the Pythagorean theorem will verify that the function h here is a continuous, subadditive interval function. The existence of the integral can be established by a repetition of the argument of Lemma 8.8.

Thus the integral

$$\int_a^b \sqrt{[dF(x)]^2 + [dG(x)]^2}$$

in the definition must necessarily be determined, although it might have an infinite value. It will have a finite value if h has bounded variation. That follows from a simple computation:

$$\max \left\{ \int_a^b |dF(x)|, \int_a^b |dG(x)| \right\} \leq \int_a^b \sqrt{[dF(x)]^2 + [dG(x)]^2}$$

and

$$\int_a^b \sqrt{[dF(x)]^2 + [dG(x)]^2} \leq \int_a^b |dF(x)| + \int_a^b |dG(x)|.$$

■

8.12.1 Formula for the length of curves

In the elementary (computational) calculus one usually assumes that a curve is given by a pair of continuously differentiable functions (i.e., a pair F, G of continuous functions for which F' and G' are also continuous). In that case the

familiar formula for length used in elementary applications is

$$\int_a^b \sqrt{[F'(x)]^2 + [G'(x)]^2} dx.$$

We study this now. Note that the formula is rather compelling if we think that $dF(x) = F'(x) dx$ and $dG(x) = G'(x) dx$ would be possible here.

Lemma 8.32 *For any pair of continuous functions $F, G : [a, b] \rightarrow \mathbb{R}$ of bounded variation on $[a, b]$ define the following function*

$$L(t) = \int_a^t \sqrt{[dF(x)]^2 + [dG(x)]^2} \quad (a < t \leq b).$$

Then

$$L'(t) = \sqrt{[F'(t)]^2 + [G'(t)]^2}$$

almost everywhere in $[a, b]$.

Proof. We are now quite familiar with the zero variation implies zero derivative a.e. rule. This is all that is needed here to establish this fact, since the statement in the Lemma can be expressed, by the Henstock zero variation criterion, as

$$\int_a^b \left| dL(x) - \sqrt{[dF(x)]^2 + [dG(x)]^2} \right| = 0.$$

Lemma 8.33 *The function L in the lemma is Vitali absolutely continuous if and only if both F and G are Vitali absolutely continuous.*

Proof. This follows easily from the inequalities of Lemma 8.31. ■

The length of the curve is now available as a familiar formula precisely in the case where the two functions defining the curve are absolutely continuous.

Lemma 8.34 For any pair of continuous functions $F, G : [a, b] \rightarrow \mathbb{R}$ of bounded variation on $[a, b]$,

$$\int_a^b \sqrt{[dF(x)]^2 + [dG(x)]^2} \geq \int_a^b \sqrt{[F'(x)]^2 + [G'(x)]^2} dx.$$

The two expressions are equal if and only if both F and G are Vitali absolutely continuous on $[a, b]$.

Proof. Using the function L introduced above we see that this assertion is easily deduced from the fact that

$$L(t) \geq \int_a^t L'(x) dx$$

with equality precisely when L is Vitali absolutely continuous. ■

Exercises

Exercise 755 For any continuous function $F : [a, b] \rightarrow \mathbb{R}$ define the length of the graph of F to mean

$$\int_a^b \sqrt{[dx]^2 + [dF(x)]^2}.$$

Show that the graph has finite length if and only if F has bounded variation. Discuss the availability of the familiar formula for length used in elementary applications:

$$\int_a^b \sqrt{1 + [F'(x)]^2} dx.$$

□

Exercise 756 Let $F, G : [a, b] \rightarrow \mathbb{R}$ where $[a, b]$ is a compact interval. Suppose that the Hellinger integral²

$$H(t) = \int_a^t \frac{dF(x) dG(x)}{dx} \quad (a < t \leq b)$$

exists. Show that $H'(t) = F'(t)G'(t)$ at almost every point t in $[a, b]$ at which both F and G are differentiable. Answer □

²Named after Ernst Hellinger (1883–1950).

Exercise 757 (Reduction theorem) Let $F, G : [a, b] \rightarrow \mathbb{R}$ where $[a, b]$ is a compact interval. Suppose that F is Vitali absolutely continuous on $[a, b]$ and that G is a Lipschitz function. Show that

$$\int_a^t \frac{dF(x) dG(x)}{dx} = \int_a^b F'(x) dG(x) = \int_a^b F'(x) G'(x) dx.$$

□

Exercise 758 Let $F, G : [a, b] \rightarrow \mathbb{R}$ where $[a, b]$ is a compact interval. Suppose that F is Vitali absolutely continuous on $[a, b]$ and that G is the indefinite integral of a function of bounded variation. Show that

$$\int_a^t \frac{dF(x) dG(x)}{dx} = \int_a^b F'(x) dG(x) = \int_a^b F'(x) G'(x) dx.$$

□

Chapter 9

Nonabsolutely Integrable Functions

The study of the Lebesgue integral in Chapter 7 usually marks the culmination of the study of integration theory on the real line for most mathematics students. They are prepared now for the more abstract theories of integration on measure spaces and studies of the important function spaces.

But the story is still not complete; part of the narrative remains. What about those functions that are integrable, but not absolutely integrable? If f is integrable on an interval $[a, b]$ but

$$\int_a^b |f(x)| dx = \infty$$

then f is not Lebesgue integrable. Its indefinite integral

$$F(x) = \int_a^x f(t) dt$$

has infinite variation on the interval $[a, b]$ since it is always true that

$$V(F, [a, b]) = \int_a^b |f(x)| dx.$$

To complete the story of the integral on the real line we must persist¹ to study the nonabsolute case and to the study of

¹**Note to the instructor:** Well *you* may not want to persist. These topics, while well-known to all specialists in real analysis, are not necessary to the backgrounds of all students, who should be encouraged now to study general measure theory and return to this subject later. The level of

indefinite integrals that do not have bounded variation. Most of the theory was developed in the decades shortly after Lebesgue's thesis. The standard account is given in

Stanislaw Saks, *Theory of the Integral*. 2nd revised edition. English translation by L. C. Young. Monografie Matematyczne, vol. 7. Warsaw, 1937.

and much of what we shall do can be found there but expressed in different language. Many mathematicians know none of this theory since the usual courses of instruction move directly to the measure-theoretic treatment of integration theory that does not address such questions.

Since we have committed our text to an account of the calculus integral we must forge ahead. The Lebesgue integral does not encompass the calculus integral for there are derivatives that are unbounded and nonabsolutely integrable. All bounded derivatives are, of course, Lebesgue integrable so that it is in the realm of the unbounded derivatives and some rather delicate considerations that this chapter will lead.

9.1 Variational Measures

The Jordan variation that we studied extensively in Chapters 3 and 8 is restricted to the study of functions of bounded variation on a compact interval $[a, b]$. When $V(f, [a, b]) = \infty$ there is not much more to be said. For a large part of the calculus program this is a sufficiently useful tool. But there are differentiable functions which do not have bounded variation and all nonabsolutely integrable functions have indefinite integrals that are not of bounded variation.

Jordan's theory was extended in the early 20th century to handle functions of finite variation on arbitrary compact sets by A. Denjoy, N. Lusin, and S. Saks. This theory was clarified later by the introduction, by R. Henstock, of measures carrying the variational information of a function. This theory includes the Jordan version and the Denjoy-Lusin-Saks versions and is the appropriate technical tool for the full range of problems arising in the calculus program.

We have already, in Chapter 8, introduced the Lebesgue-Stieltjes measures λ_f and we return to that study now with an additional variational measure that is dual to the measure λ_f called the fine variation.

this chapter is, accordingly, somewhat raised above the expository level of the preceding chapters.

9.1.1 Full and fine variational measures

The variation of a function f on an interval $[a, b]$ is described by the identity

$$V(f, [a, b]) = \sup_{\pi} \left(\sum_{([u, v], w) \in \pi} |f(v) - f(u)| \right) \quad (9.1)$$

where the supremum is taken over all possible partitions π of the interval $[a, b]$. We recall that a similar expression describes the Lebesgue-Stieltjes measure

$$\lambda_f(E) = \inf_{\beta} \left\{ \sup_{\pi \subset \beta} \left(\sum_{([u, v], w) \in \pi} |f(v) - f(u)| \right) \right\} \quad (9.2)$$

where the supremum is taken over all possible subpartitions π contained in β and the infimum is taken over all full covers β of the set E . The two expressions (9.1) and (9.2) are clearly closely related but the exact relationship needs some thinking (see Exercise 770).

The generalization of Lebesgue measure to the Lebesgue-Stieltjes measure arises by replacing $(v - u)$ by $|f(v) - f(u)|$. It is more convenient for our purposes to write

$$\Delta f([u, v]) = f(v) - f(u)$$

so that $\Delta f(I)$ is an interval function that computes the *increment* of the function f on the interval I . This is often useful in conjunction with the notation $\omega f(I)$ denoting the *oscillation* of the function f on the interval I , defined, we recall, as

$$\omega f(I) = \sup_{u, v \in I} |f(v) - f(u)|.$$

We review the Lebesgue-Stieltjes measure construction and add to it a new variational measure based on fine covers instead of full covers.

Definition 9.1 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function and β a covering relation. We write

$$V(\Delta f, \beta) = \sup_{\pi \subset \beta} \left\{ \sum_{([u, v], w) \in \pi} |\Delta f([u, v])| \right\}$$

where the supremum is taken over all subpartitions π contained in β .

Definition 9.2 (Full and Fine Variations) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and let E be any set of real numbers. Then we define the full and fine variational measures associated with f by the expressions:

$$\lambda_f(E) = V^*(\Delta f, E) = \inf\{V(\Delta f, \beta) : \beta \text{ a full cover of } E\}$$

and

$$\lambda_f^*(E) = V_*(\Delta f, E) = \inf\{V(\Delta f, \beta) : \beta \text{ a fine cover of } E\}.$$

Note that the star $\star$ (not an asterisk $*$) indicates the fine variation. In general the inequality $\lambda_f^*(E) \leq \lambda_f(E)$ holds and identity holds only for a certain (important) class of functions. These set functions share the same properties as the measure λ . Specifically they are countably subadditive for sequences of sets and they are countably additive for disjoint sequences of closed sets.

9.1.2 Finite variation and σ -finite variation

Definition 9.2 allows us to extend the notion of bounded variation to describe the situation on arbitrary sets.

1. f has *bounded variation* on an interval $[a, b]$ if $V(f, [a, b]) < \infty$.
2. f has *finite variation* on a set E if $\lambda_f(E) < \infty$.
3. f has *σ -finite variation* on a set E if there is a sequence of sets $\{E_n\}$ covering E and $\lambda_f(E_n) < \infty$ for each $n = 1, 2, 3, \dots$.

We shall state now and prove (eventually) that the Lebesgue differentiation theorem of Chapter 5 can be extended to this larger class of functions. Recall that our original statement required that the function have bounded variation on the whole of some interval.

Theorem 9.3 (Lebesgue differentiation theorem) Let f be a continuous function defined on some open set that contains a set E on which f has σ -finite variation. Then f is differentiable λ -almost everywhere in E and has a finite or infinite derivative λ_f -almost everywhere in E .

The proof follows from Theorem 9.20 that we shall prove much later.

9.1.3 The Vitali property

The two measures λ_f and λ_f^* together express the variation of the function f . We recall that they are analogous to the full and fine versions of Lebesgue measure, λ^* and λ_* . Those two measures are identical because of the Vitali covering theorem and the identity

$$\lambda_f = \lambda_f^*$$

(when it holds) would be considered a generalization of the Vitali covering theorem. It is not the case that $\lambda_f = \lambda_f^*$ in general, but for a most important class of functions this will be true. When the Vitali theorem holds for these measures we say that the function f has the Vitali property.

Definition 9.4 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and let E be any set of real numbers. Then we say that the function f has the Vitali property on E provided that the two measures λ_f and λ_f^* agree on all subsets of E .*

9.1.4 Kolmogorov equivalence

The variation describes a convenient equivalence relation between functions. The notion originated with the Russian mathematician Kolmogorov, and was exploited in this context by Henstock who used the terminology “variational equivalence.”

Definition 9.5 (Kolmogorov equivalent) *Two functions f and g are said to be Kolmogorov equivalent on E if*

$$V^*(\Delta f - \Delta g, E) = 0.$$

By means of this equivalence relation we can lift a number of properties that we already know for functions of bounded variation to a more general class of functions. When two functions are equivalent in this sense then they must share many properties in common. Here is a list of such properties. Proofs are left for the exercises.

Implications of Kolmogorov equivalence. If the functions f and g are Kolmogorov equivalent on E then:

1. $f'(x) = g'(x)$ at almost every point in E at which g is differentiable. [A partial converse is given in Exercise 764.]

2. f is continuous at every point in E at which g is continuous.
3. f is locally bounded at every point in E at which g is locally bounded.
4. f has the Vitali property on E if and only if g has the Vitali property on E .
5. f has finite variation on E if and only if g has finite variation on E .
6. f has zero variation on E if and only if g has zero variation on E .
7. $\lambda_f(E) = \lambda_g(E)$ and $\lambda_f^*(E) = \lambda_g^*(E)$.

9.1.5 Variation of continuous, increasing functions

In special cases it is easy to estimate the full and fine variations. Note that as a result of this first computation we see that continuous, increasing functions possess the Vitali property.

Theorem 9.6 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be continuous and strictly increasing. Then, for any set E ,*

$$\lambda_f^*(E) = \lambda_f(E) = \lambda(f(E))$$

and f has the Vitali property on every set.

Proof. If β is a full [fine] cover of E then check that

$$\beta' = \{(f(I), f(x)) : (I, x) \in \beta\}$$

is a full [fine] cover of $f(E)$. Note too that $\Delta f(I) = \lambda(f(I))$ for such a function. From this we deduce that

$$\lambda^*(f(E)) = \lambda_f(E)$$

and

$$\lambda_*(f(E)) = \lambda_f^*(E).$$

By the Vitali covering theorem $\lambda^* = \lambda_*$ so that the identity in the theorem now follows. ■

9.1.6 Variation and image measure

In general the full variation is larger than the image measure.

Theorem 9.7 For an arbitrary function $f : \mathbb{R} \rightarrow \mathbb{R}$ and any real set E ,

$$\lambda(f(E)) \leq \lambda_f(E).$$

Proof. Let $\lambda_f(E) < t$ and select a full cover β of E so that $V(\Delta f, \beta) < t$. We apply the decomposition lemma, Lemma 5.6, for β . There is an increasing sequence of sets $\{E_n\}$ with $E = \bigcup_{n=1}^{\infty} E_n$ and a sequence of nonoverlapping compact intervals $\{I_{kn}\}$ covering E so that if x is any point in E_n and I is any subinterval of I_k that contains x then (I, x) belongs to $\beta([E_n \cap I_{kn}])$.

Thus let us estimate the λ -measure of the set $f(E_n \cap I_{kn})$. Our estimate need only be crude: if $f(x_1), f(x_2)$ with $x_1 < x_2$ are any two points in this set then certainly $([x_1, x_2], x_1) \in \beta(I_k)$. Thus

$$|f(x_1) - f(x_2)| = |\Delta f([x_1, x_2])| \leq V(\Delta f, \beta(I_{kn}))$$

so it follows that

$$\lambda(f(E_n \cap I_{kn})) \leq V(\Delta f, \beta(I_{kn})).$$

Hence, using Exercise 759 and usual properties of Lebesgue measure,, we have that

$$\lambda(f(E_n)) \leq \sum_k \lambda(f(E_n \cap I_{kn})) \leq \sum_k V(\Delta f, \beta(I_{kn})) \leq V(\Delta f, \beta) < t.$$

Note that the sequence $\{E_n\}$ is expanding and that its union is the whole set E ; it follows that $\{f(E_n)\}$ is expanding and that its union is the whole set $f(E)$. Accordingly then, by Theorem 7.14,

$$\lim_{n \rightarrow \infty} \lambda(f(E_n)) = \lambda(f(E)).$$

It follows that

$$\lambda(f(E)) \leq t.$$

Since t was merely chosen so that $\lambda_f(E) < t$ it follows that $\lambda(f(E)) \leq \lambda_f(E)$ as required. ■

9.1.7 Variational classifications of real functions

Let us review and enlarge some of our terminology for the behavior of functions. All of the following ideas are expressible in the language of the variation. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and let E be any set of reals.

(zero variation) f has *zero variation* on E if $\lambda_f(E) = 0$.

(finite variation) f has *finite variation* on E if $\lambda_f(E) < \infty$.

(σ -finite variation) f has *σ -finite variation* on E if $E \subset \bigcup_{k=1}^{\infty} E_k$ so that $\lambda_f(E_k) < \infty$ for each $k = 1, 2, 3, \dots$

(Kolmogorov equivalent) f and g are *Kolmogorov equivalent* on E if $V^*(\Delta f - \Delta g, E) = 0$.

(Vitali property on a set) f has the *Vitali property* on E provided that, for all subsets A of E , $\lambda_f(A) = \lambda_f^*(A)$.

(continuous at a point) f is *continuous* at a point x_0 provided that $\lambda_f(\{x_0\}) = 0$.

(weakly continuous at a point) f is *weakly continuous* at a point x_0 provided that $\lambda_f^*(\{x_0\}) = 0$.

(λ -absolutely continuous on a set) f is *λ -absolutely continuous*² on E provided that, for every set $N \subset E$ that has Lebesgue measure zero, $\lambda_f(N) = 0$.

(λ -singular on E) f is *λ -singular* on E provided $\lambda_f(E \setminus N) = 0$ for some set $N \subset E$ that has Lebesgue measure zero.

(mutually singular) Two functions f and g are said to be *mutually singular* on a set E if $E = E_1 \cup E_2$ and $\lambda_f(E_2) = \lambda_g(E_1) = 0$.

(saltus function) f is a *saltus function* on an open interval (a, b) if there is a countable set C so that $\lambda_f((a, b) \setminus C) = 0$ and $\lambda_f((a, b) \cap C) < \infty$.

Since each of these terms is definable or describable directly in terms of the variational measures it should be expected that there are many interrelationships. Some of these are explored in the exercises.

²We previously referred to this simply as *absolutely continuous* without specifying the measure to which λ_f is being compared.

Exercises

Exercise 759 Let β be a covering relation and $f : \mathbb{R} \rightarrow \mathbb{R}$. If $\{I_k\}$ is a sequence of nonoverlapping subintervals of an interval I (open or closed) then show that

$$\sum_{k=1}^{\infty} V(\Delta f, \beta(I_k)) \leq V(\Delta f, \beta(I)).$$

□

Exercise 760 (Subadditivity property) Let h_1 and h_2 be real-valued functions defined on interval-point pairs. Then, for any set E , show that

$$V_*(h_1 + h_2, E) \leq V_*(h_1, E) + V^*(h_2, E)$$

and

$$V^*(h_1 + h_2, E) \leq V^*(h_1, E) + V^*(h_2, E).$$

Answer □

Exercise 761 Let $f, g : \mathbb{R} \rightarrow \mathbb{R}$. Write $f \sim g$ on E if f and g are Kolmogorov equivalent on E . Show that this is an equivalence relation. □

Exercise 762 Let $f, g : \mathbb{R} \rightarrow \mathbb{R}$. Show that, if f and g are Kolmogorov equivalent on a set E , then $\lambda_f(E) = \lambda_g(E)$ and $\lambda_f^*(E) = \lambda_g^*(E)$. □

Exercise 763 Let $f, g : \mathbb{R} \rightarrow \mathbb{R}$. Show that, if f and g are Kolmogorov equivalent on each of the sets $E_1, E_2, E_3, \dots$ then f and g are Kolmogorov equivalent on the union of these sets. □

Exercise 764 Let $f, g : \mathbb{R} \rightarrow \mathbb{R}$. Show that, if $f'(x) = g'(x)$ at every point of a set E then f and g are Kolmogorov equivalent on E . □

Exercise 765 Let $f : \mathbb{R} \rightarrow \mathbb{R}$. Show that f is λ -singular on a set E if $f'(x) = 0$ at almost every point x of E . Answer □

Exercise 766 Let $f : \mathbb{R} \rightarrow \mathbb{R}$. Show, conversely, that if f is λ -singular on a set E then $f'(x) = 0$ at almost every point x of E . □

Exercise 767 Show that if $f : \mathbb{R} \rightarrow \mathbb{R}$ has finite variation or σ -finite variation on a set E then f is continuous at each point of E with countably many exceptions. Answer $\square$

Exercise 768 Show that a function $f : \mathbb{R} \rightarrow \mathbb{R}$ is weakly continuous at a point x_0 if and only if there are sequences $c_n \nearrow x_0$ and $d_n \searrow x_0$ so that $d_n - c_n > 0$ and

$$f(d_n) - f(c_n) \rightarrow 0.$$

Answer $\square$

Exercise 769 Let $f : \mathbb{R} \rightarrow \mathbb{R}$. Show that f must be weakly continuous at every point with at most countably many exceptions. Answer $\square$

Exercise 770 Let $f : \mathbb{R} \rightarrow \mathbb{R}$. Establish the following relation between the Jordan variation and the variational measures:

$$\lambda_f((a, b)) \leq V(f, [a, b]) \leq \lambda_f([a, b]) = \lambda_f((a, b)) + \lambda_f(\{a\}) + \lambda_f(\{b\}).$$

In particular show that

$$\lambda_f((a, b)) = \lambda_f([a, b]) = V(f, [a, b])$$

if f is continuous at a and b . $\square$

Exercise 771 Let $f : \mathbb{R} \rightarrow \mathbb{R}$. Show that f has bounded variation on $[a, b]$ if and only if f has finite variation on (a, b) . Give an example to show that, even so, $V(f, [a, b])$ may be different from $\lambda_f((a, b))$. Answer $\square$

Exercise 772 Let $E \subset (a, b)$ be a compact set and let $\{(a_i, b_i)\}$ be the component intervals of $(a, b) \setminus E$. Suppose that f is a continuous function satisfying $f(x) = 0$ for all $x \in E$ and that

$$\sum_i \omega f([a_i, b_i]) < \infty.$$

Show that $\lambda_f(E) = 0$. $\square$

Exercise 773 (local recurrence) A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is locally recurrent at a point x if there is a sequence of points x_n with $x_n \neq x$ and $\lim_{n \rightarrow \infty} x_n = x$ so that $f(x) = f(x_n)$ for all n . Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and suppose that f is locally recurrent at every point of a set E . Show that $\lambda_f^*(E) = 0$. Answer $\square$

Exercise 774 (local monotonicity) A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is locally nondecreasing at a point x if there is a $\delta > 0$ so that $\Delta f(I) \geq 0$ for every compact interval I containing x for which $\lambda(I) < \delta$. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and suppose that f is locally nondecreasing at every point of a set E and that $\lambda_f(\{x\}) < \infty$ for each x in E . Show that f has σ -finite variation on E . Answer $\square$

Exercise 775 (continuous functions have σ -finite fine variation) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function. Show that λ_f^* must be σ -finite. Answer $\square$

Exercise 776 (Lebesgue differentiation theorem) Prove Theorem 9.3:

Let f be a continuous function defined on some open set that contains a set E on which f has σ -finite variation. Then f is differentiable at almost every point of E .

Hint: You may assume here the conclusion of Theorem 9.20 that there is a sequence of compact sets covering E on each of which f is Kolmogorov equivalent to some continuous function of bounded variation. Answer $\square$

9.2 Derivates and variation

If the derivates of a function $f : \mathbb{R} \rightarrow \mathbb{R}$ are finite on a set E this has implications for the variation λ_f on E .

9.2.1 Ordinary derivates and variation

Theorem 9.8 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and suppose that f is differentiable at every point x of a set E . Then

$$\lambda_f(E) = \lambda_f^*(E) = \int_E |f'(x)| dx.$$

In particular f has σ -finite variation, is λ -absolutely continuous, and has the Vitali property on that set.

Proof. The fact that $f'(x)$ exists on E leads immediately to the variational identity

$$V^*(\Delta f - f'\lambda, E) = 0.$$

From this, using Exercise 760, we can deduce that

$$V^*(\Delta f, E) \leq V^*(\Delta f - f'\lambda, E) + V^*(f'\lambda, E)$$

and hence that

$$\lambda_f(E) = V^*(\Delta f, E) \leq V^*(f'\lambda, E) = \int_E |f'(x)| dx.$$

The opposite inequality is proved the same way.

Again, using the other inequality in Exercise subaddprop, we can deduce that

$$V_*(f'\lambda, E) \leq V^*(\Delta f - f'\lambda, E) + V_*(\Delta f, E) = \lambda_f^*(E)$$

Since f' is measurable the identity

$$\int_E |f'(x)| dx = V_*(f'\lambda, E) = V^*(f'\lambda, E)$$

can be used to complete the proof. ■

Theorem 9.9 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and suppose at every point x of a set E that $\lambda_f(\{x\}) < \infty$ and that either $\overline{D}f(x) < \infty$ or $\underline{D}f(x) > -\infty$. Then f has σ -finite variation in E .*

Proof. For example let us consider that the set E consists of all points at which $\underline{D}f(x) > -\infty$. Write

$$E_n = \{x : \underline{D}f(x) > -n\}.$$

Note that E is the union of the sequence of sets $\{E_n\}$.

Observe that the function $f_n(x) = f(x) + nx$ is locally nondecreasing at each $x \in E_n$. It follows (from Exercise 774) that f_n has σ -finite variation on E_n . But

$$\lambda_f \leq \lambda_{f_n} + n\lambda.$$

Thus f too has σ -finite variation on E_n . In consequence, f has σ -finite variation on E . ■

9.2.2 Dini derivatives and variation

For many functions a closer analysis is needed than would be available using the upper and lower derivatives: we require one-sided versions.

Definition 9.10 (Dini derivatives) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and suppose that $x \in \mathbb{R}$. Then the four values

$$D^+ f(x) = \inf_{\delta > 0} \sup \left\{ \frac{f(x+h) - f(x)}{h} : 0 < h < \delta \right\}$$

$$D_+ f(x) = \sup_{\delta > 0} \inf \left\{ \frac{f(x+h) - f(x)}{h} : 0 < h < \delta \right\}$$

$$D^- f(x) = \inf_{\delta > 0} \sup \left\{ \frac{f(x) - f(x-h)}{h} : 0 < h < \delta \right\}$$

$$D_- f(x) = \sup_{\delta > 0} \inf \left\{ \frac{f(x) - f(x-h)}{h} : 0 < h < \delta \right\}$$

are called the Dini derivatives of f at x .

We do not need much more information than this for our main theorem. The reader interested in pursuing the Dini derivatives further should try Exercises 777–786. We will return in Section 9.14 to the Dini derivatives and show how a continuous function can be recovered by integrating one of its Dini derivatives.

Theorem 9.11 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function and suppose that at every point x of a set E either

$$-\infty < D_+ f(x) \leq D^+ f(x) < \infty$$

or

$$-\infty < D_- f(x) \leq D^- f(x) < \infty.$$

Then f has σ -finite variation in E and is λ -absolutely continuous there.

Proof. We first show that, for any positive integer c , f has σ -finite variation and is λ -absolutely continuous on the set of points

$$A = \{x : -c < D_+ f(x) \leq D^+ f(x) < c\}.$$

The geometry of this situation is expressed by the covering relation

$$\beta = \{[x, x+h], x) : |\Delta f([x, x+h])| < c\lambda([x, x+h])\}.$$

This relation has none of the properties we have so far encountered, but a modification of our methods will handle.

First apply the ideas of the decomposition from Section 5.6 for β . There is an increasing sequence of sets $\{A_n\}$ with $A = \bigcup_{n=1}^{\infty} A_n$ and a sequence of compact intervals $\{I_{kn}\}$ covering A so that if x is any point in A_n and $[x, x+h]$ is any subinterval of I_{kn} then $([x, x+h], x)$ belongs to β .

In particular if $\{[c_i, d_i]\}$ is a sequence of subintervals of I_{kn} with endpoints in the set A_n , then a brief computation shows that

$$\sum_{i=1}^{\infty} \omega f([c_i, d_i]) \leq \sum_{i=1}^{\infty} 2c\lambda([c_i, d_i]) \leq 2c\lambda(I_{kn}).$$

Let C_{nk} denote the closure of the set $A_n \cap I_{kn}$. Since f is continuous this same inequality extends to points in that closure. Thus if $\{[c_i, d_i]\}$ is a sequence of intervals with endpoints in the compact set C_{nk} , then

$$\sum_{i=1}^{\infty} \omega f([c_i, d_i]) \leq \sum_{i=1}^{\infty} 2c\lambda([c_i, d_i]) \leq 2c\lambda(I_{kn}) < \infty.$$

Define a function g_n so that $g_n(x) = f(x)$ for all $x \in C_{nk}$ and extend to all of the real line so as to be continuous and linear on all of the complementary intervals to C_{nk} . Such a function g_n is evidently continuous and has bounded variation. The same inequality shows that g_n is absolutely continuous in the sense of Vitali and so also λ -absolutely continuous.

The computations of Exercise 772 can be used here to check that

$$V^*(\Delta f - \Delta g_n, C_{kn}) = 0.$$

This shows that f is Kolmogorov equivalent on each set C_{nk} to a continuous function of bounded variation. In particular λ_f is finite on each set C_{nk} . It follows that λ_f is σ -finite on A . The function f also inherits from g_n the property of being λ -absolutely continuous on C_{nk} .

Finally the set E of the theorem can be expressed as a union of a sequence of sets of the same type as A , so that λ_f is σ -finite and vanishes on null subsets of each member of the sequence. The theorem follows. ■

9.2.3 Lipschitz numbers

A Lipschitz condition on a function is a global upper estimate of the ratio

$$\left| \frac{F(y) - F(x)}{y - x} \right| = \left| \frac{\Delta F([x, y])}{\lambda([x, y])} \right|$$

We can make this same estimate locally in which case the estimates are called Lipschitz numbers and they serve as a local estimate of the growth of a function. We refine this a bit by introducing a lower estimate as well. In Section 9.2.4 we show how these numbers relate to the variations.

If $h(I, x)$ is any function which assigns real values to interval-point pairs we recall that in Section 5.6.2 we introduced the following notation for the limits:

$$\limsup_{(I,x) \Rightarrow x} = \inf_{\delta > 0} (\sup\{h(I, x) : \lambda(I) < \delta, x \in I\})$$

and

$$\liminf_{(I,x) \Rightarrow x} = \sup_{\delta > 0} (\inf\{h(I, x) : \lambda(I) < \delta, x \in I\}).$$

These are just convenient expressions for the lower and upper limits of $h(I, x)$ as the interval I (always assumed to contain x) shrinks to the point x . As usual if the $\limsup$ and $\liminf$ are same then the common value (including ∞ and $-\infty$) would be written as

$$\lim_{(I,x) \Rightarrow x} h(I, x).$$

When working with such limits Exercises 793 and 794 offer useful estimates of some associated variations.

Definition 9.12 Let $f : \mathbb{R} \rightarrow \mathbb{R}$. Then

$$\overline{\text{lip}}_f(x) = \limsup_{(I,x) \Rightarrow x} \left| \frac{\Delta f(I)}{\lambda(I)} \right|$$

$$\underline{\text{lip}}_f(x) = \liminf_{(I,x) \Rightarrow x} \left| \frac{\Delta f(I)}{\lambda(I)} \right|$$

are called the upper and lower Lipschitz number of f at a point x .

Lemma 9.13 Let $f : \mathbb{R} \rightarrow \mathbb{R}$. For any real number r the sets

$$\{x : \overline{\text{lip}}_f(x) < r\} \text{ and } \{x : \underline{\text{lip}}_f(x) < r\}$$

are measurable.

This is nearly identical to Lemma 7.18.

9.2.4 Six growth lemmas

The growth lemmas we present all follow easily from the general limit lemmas of Exercises 793 and 794. Proofs are left to the student. They can be considered as generalizations of these simple facts:

1. If $f'(x) \leq r$ for all x then

$$f(b) - f(a) \leq r(b - a).$$

2. If $f'(x) \geq r$ for all x then

$$f(b) - f(a) \geq r(b - a).$$

Now, however, the derivative is replaced by upper and lower Lipschitz estimates, the interval $[a, b]$ is replaced by an arbitrary set and the increments are replaced by variational measures.

Lemma 9.14 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$. If $\underline{\text{lip}}_f(z) < r$ for every $z \in E$ then*

$$\lambda_f^*(E) \leq r\lambda(E).$$

Lemma 9.15 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$. If $\overline{\text{lip}}_f(z) > r > 0$ for every $z \in E$ then*

$$r\lambda(E) \leq \lambda_f(E).$$

Lemma 9.16 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$. If $\overline{\text{lip}}_f(z) < r$ for every $z \in E$ then*

$$\lambda_f(E) \leq r\lambda(E).$$

Lemma 9.17 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$. If $\underline{\text{lip}}_f(z) > r > 0$ for every $z \in E$ then*

$$r\lambda(E) \leq \lambda_f^*(E).$$

Lemma 9.18 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$. If $\lambda_f(E) < \infty$, then $\overline{\text{lip}}_f(x) < \infty$ for almost every point x in E .*

Lemma 9.19 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$. If $\lambda_f^*(E) < \infty$ then $\underline{\text{lip}}_f(z) < \infty$ for almost every point x in E .*

Exercises

Exercise 777 Show that

$$\underline{D}f(x) \leq D_+f(x) \leq D^+f(x) \leq \overline{D}f(x) \text{ and } \overline{D}f(x) = \max\{D^-f(x), D^+f(x)\}.$$

□

Exercise 778 (Grace Chisolm Young) Let $f : \mathbb{R} \rightarrow \mathbb{R}$. Show that the sets of points

$$\{x : D^-f(x) < D_+f(x)\}$$

and

$$\{x : D^+f(x) < D_-f(x)\}$$

are both countable.

Answer □

Exercise 779 (Beppo Levi) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and suppose that f has one-sided derivatives $f'_+(x)$ and $f'_-(x)$ at each point of a set E . Show that the set of points x in E at which

$$f'_+(x) \neq f'_-(x)$$

is countable.

□

Exercise 780 It is easy to misinterpret the theorem of Beppo Levi (Exercise 779). To avoid this construct a continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$ so that for some uncountable set E the right-hand derivative $f'_+(x)$ exists at each point of E and the left-hand derivative $f'_-(x)$ fails to exist at each point of E .

□

Exercise 781 (William Henry Young) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function. Show that the sets of points

$$\{x : D^-f(x) = D^+f(x)\}$$

and

$$\{x : D_-f(x) = D_+f(x)\}$$

are both residual subsets of $\mathbb{R}$.

Answer □

Exercise 782 Let $f : [a, b] \rightarrow \mathbb{R}$ be a continuous function. Show that the set of points at which f has a right-hand derivative but no left-hand derivative is a meager subset of $[a, b]$. $\square$

Exercise 783 Let $f : [a, b] \rightarrow \mathbb{R}$ be a continuous function with $f([a, b]) = [c, d]$. Write

$$D = \{x \in [a, b] : D^+ f(x) \leq 0\}.$$

Show that either f is nondecreasing on $[a, b]$ or else $f(D)$ contains a compact subinterval of $[c, d]$. Answer $\square$

Exercise 784 (Anthony P. Morse) Let $f : [a, b] \rightarrow \mathbb{R}$ be a continuous function with $f([a, b]) = [c, d]$. Write

$$A = \{x \in [a, b] : D^+ f(x) \geq 0\},$$

$$B = \{x \in [a, b] : D^+ f(x) < 0\},$$

and

$$C = \{x \in [a, b] : D^+ f(x) = 0\}.$$

Suppose that A is dense in $[a, b]$. Show that B is a meager subset of $[a, b]$ and $f(B)$ is a meager subset of $[c, d]$. Moreover, show that either f is nondecreasing on $[a, b]$ or else $f(C)$ contains a residual subset of some compact subinterval of $[c, d]$. Answer $\square$

Exercise 785 (Darboux property of Dini derivatives) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function and suppose that the Dini derivative $D^+ f(x)$ is unbounded both above and below on each interval. Show, for every real number r and compact interval $[a, b]$, that f maps the set

$$E_r = \{x \in [a, b] : D^+ f(x) = r\}$$

onto a residual subset of some compact interval. (In particular $D^+ f(x)$ assumes every real number at many points in any subinterval.) $\square$

Exercise 786 For any continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$ and any real number r show that the sets

$$\{x : D^+ f(x) \leq r\} \text{ and } \{x : D_+ f(x) \leq r\}$$

are Lebesgue measurable. Answer $\square$

9.2. DERIVATES AND VARIATION

367

Exercise 787 Let $f : \mathbb{R} \rightarrow \mathbb{R}$. Verify that

$$\overline{\text{lip}}_f(x) = \max\{|\overline{D}f(x)|, |\underline{D}f(x)|\}$$

and also

$$\overline{\text{lip}}_f(x) = \max\{|\overline{D}^+ f(x)|, |\underline{D}^+ f(x)|, |\overline{D}^- f(x)|, |\underline{D}^- f(x)|\}.$$

□

Exercise 788 Let $f : \mathbb{R} \rightarrow \mathbb{R}$. Suppose that f has a derivative at x (finite or infinite). Show that $\overline{\text{lip}}_f(x) = \underline{\text{lip}}_f(x) = |f'(x)|$.

□

Exercise 789 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function, and suppose that $\overline{\text{lip}}_f(x) = \underline{\text{lip}}_f(x) < \infty$. Show that f has a finite derivative at x and that

$$\overline{\text{lip}}_f(x) = \underline{\text{lip}}_f(x) = |f'(x)|.$$

Answer □

Exercise 790 If $f : \mathbb{R} \rightarrow \mathbb{R}$ is continuous and $\underline{\text{lip}}_f(x) = \infty$ show that either $f'(x) = \infty$ or $f'(x) = -\infty$. Give an example to show that continuity cannot be dropped.

□

Exercise 791 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function. Show that $\underline{\text{lip}}_f(x) < \infty$ for almost every point x .

Answer □

Exercise 792 For this exercise and the next three exercises we shall use the following generalized variations. Let h be any real-valued function defined on interval-point pairs and define

$$V(h, \beta) = \sup_{\pi \subset \beta} \left\{ \sum_{([u,v], w) \in \pi} |h([u,v], w)| \right\}$$

where the supremum is taken over all π , arbitrary subpartitions contained in β ;

$$h^*(E) = \inf\{V(h, \beta) : \beta \text{ a full cover of } E\}$$

and

$$h_*(E) = \inf\{h, \beta) : \beta \text{ a fine cover of } E\}.$$

Show that h^* and h_* are measures and that $h_* \leq h^*$.

□

Exercise 793 (limsup comparison lemma) Suppose that, for every x in a set E ,

$$s < \limsup_{(I,x) \Rightarrow x} \left| \frac{h(I,x)}{\lambda(I)} \right| < r$$

Show that

$$s\lambda(E) \leq V^*(h, E) \leq r\lambda(E)$$

and

$$V_*(h, E) \leq r\lambda(E).$$

Answer ☐

Exercise 794 (liminf comparison lemma) Suppose that, for every x in a set E ,

$$s < \liminf_{(I,x) \Rightarrow x} \left| \frac{h(I,x)}{k(I,x)} \right| < r$$

Show that

$$s\lambda(E) \leq V_*(h, E) \leq r\lambda(E)$$

and

$$s\lambda(E) \leq V^*(h, E).$$

Answer ☐

Exercise 795 Deduce all of the growth lemmas in Section 9.2.4 from the liminf comparison and limsup comparison lemmas (i.e., Exercises 793 and 794). ☐

Exercise 796 Let $f : \mathbb{R} \rightarrow \mathbb{R}$. If $\overline{\text{lip}}_f(z) < \infty$ for every $z \in E$ then show that f has σ -finite variation in E and is λ -absolutely continuous there. ☐

Answer ☐

9.3 Continuous functions with σ -finite variation

We begin now a deeper analysis of those continuous functions that have σ -finite full variation on a set. Because of part (3) of this theorem we now can deduce the Lebesgue differentiation theorem (Theorem 9.3) asserting that these functions are almost everywhere differentiable.

Theorem 9.20 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function and E a real set. Then the following are equivalent:*

1. f has σ -finite variation on E ,
2. there is a sequence $\{E_n\}$ of compact sets covering E so that f has finite variation on each E ,
3. there is a sequence $\{E_n\}$ of compact sets covering E so that on each E_n , f is Kolmogorov equivalent to some continuous function of bounded variation.

Proof. The implication (2) $\implies$ (1) is trivial. The implication (3) $\implies$ (2) is easy: if (3) holds then, for some continuous function of bounded variation $g_n : \mathbb{R} \rightarrow \mathbb{R}$, the equivalence relation

$$V^*(\Delta f - \Delta g_n, E_n) = 0$$

implies that $\lambda_f(E_n) = \lambda_{g_n}^*(E_n) < \infty$.

Thus the proof is completed by showing that (1) $\implies$ (3). It is enough to consider the situation for which E is a bounded set for which $\lambda_f(E) < \infty$. Choose a full cover β of E and a real number t so that

$$V(\Delta f, \beta) < t < \infty.$$

Apply the decomposition in Lemma 5.6 to β . Accordingly there is an increasing sequence of sets $\{B_n\}$ with $E = \bigcup_{n=1}^{\infty} B_n$ and a sequence of nonoverlapping compact intervals $\{I_{kn}\}$ covering E so that if x is any point in B_n and I is any subinterval of I_{kn} that contains x then (I, x) belongs to β .

Let $A_{kn} = B_n \cap I_{kn}$. We check some facts about the variation of f on A_{kn} . Suppose that $\{[a_i, b_i]\}$ is any disjointed sequence of compact subintervals of I_{kn} each of which contains at least one point, say x_i , of B_n . Then $\{([a_i, b_i], x_i)\}$ must

form a subpartition contained in β . Consequently

$$\sum_i |f(b_i) - f(a_i)| \leq V(\Delta f, \beta) < t.$$

Now let C_{kn} denote the closure of A_{kn} , i.e., C_{kn} is the smallest compact set that contains A_{kn} . We extend these considerations to estimating the variation of f on the larger set C_{kn} . Suppose now that $\{[a_i, b_i]\}$ is any disjointed sequence of compact subintervals of I_{kn} each of which contains at least one point of C_{nk} . We enlarge each interval slightly as needed to ensure that the intervals remain disjointed but contain also a point, now, of the dense subset A_{kn} . As f is continuous we can do this without much of an increase in the sums, and so we can certainly guarantee that for the given sequence $\{[a_i, b_i]\}$ that

$$\sum_i |f(b_i) - f(a_i)| < 2t < \infty.$$

Let us define a function g_{nk} so as to be equal to $f(x)$ on the compact set C_{kn} and extended to the real line so as to be linear and continuous on the intervals complementary to C_{kn} . Such a function g_{nk} is continuous and has bounded variation.

The computations of Exercise 772 can be used here to check that

$$V^*(\Delta f - \Delta g_{nk}, C_{kn}) = 0.$$

As every compact set from the sequence $\{C_{kn}\}$ can be treated the same way, we have verified the implication $(1) \implies (3)$ provided we merely relabel the full collection $\{C_{kn}\}$ as a single sequence $\{E_n\}$. ■

9.3.1 Variation on compact sets

We can refine our analysis of σ -finite variation with a few further steps.

Theorem 9.21 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function and E a compact set. Then the following are equivalent:*

1. f has σ -finite variation on E .
2. Every nonempty compact subset S of E has a portion $S \cap (a, b)$ on which f has finite variation.
3. f has σ -finite variation on every null set $Z \subset E$ that is a G_δ set.

Proof. By a G_δ set we mean a set Z of the form $Z = \bigcap_{n=1}^{\infty} G_n$ for some sequence $\{G_n\}$ of open sets. Every closed set can be written in this form.

We begin with (a) $\implies$ (b). As we have seen in Theorem 9.20, if f has σ -finite variation on E , then there is a sequence of compact sets $\{E_n\}$ covering the compact set S so that $\lambda_f(E_n) < \infty$ for each n . By the Baire category theorem there must be a portion $S \cap (a, b)$ of E contained in one at least from the sequence $\{E_n\}$. In particular, for some n , $\lambda_f(S \cap (a, b)) \leq \lambda_f(E_n) < \infty$ as required to prove (b).

Let us now prove that (b) $\implies$ (a) Suppose that every nonempty closed subset S of E has a portion $S \cap (a, b)$ on which f has finite variation. Let G denote the real set consisting of all real x with the property that there is a $\delta(x) > 0$ so that f has σ -finite variation on the set $E \cap (x - \delta(x), x + \delta(x))$. Note that

$$G = \bigcup_{x \in G} (x - \delta(x), x + \delta(x))$$

so G is open.

Consider the set $G \cap E$. Any point in this set would be contained in an open interval (c, d) with rational endpoints so that f has σ -finite variation on $G \cap (c, d)$. It follows that f has σ -finite variation on $G \cap E$. If $G \supset E$ then, we deduce that f has σ -finite variation on E as we wished to prove to verify (a).

Suppose, in order to obtain a contradiction that G does not contain E . Let $E' = E \setminus G$. This would be a nonempty closed subset of E and so, by hypothesis, there would have to be a portion $E' \cap (a, b)$ on which f has finite variation. But if f has finite variation on $E' \cap (a, b)$ and also, evidently, has σ -finite variation on $E \setminus E'$ then f must have σ -finite variation on $E \cap (a, b)$. Every point of this set should belong to G which is impossible in view of the assumption that $E' \cap (a, b)$ is a portion. This contradiction completes our proof that (b) $\implies$ (a).

The implication (a) $\implies$ (c) is trivial. To complete the proof, then, it will suffice to verify that (c) $\implies$ (b). Suppose

that f has σ -finite variation on every set $Z \subset E$ that is a $\mathcal{G}_\delta$ set of λ -measure zero.

Let S be a nonempty closed subset of E . To verify (b) we need to find a portion of S on which f has finite variation. If S is a null set then we are almost there. A closed set is also of type $\mathcal{G}_\delta$. Thus, λ_f is σ -finite on S by hypothesis. As we have already argued above, in this situation we are assured that S has a portion $S \cap (a, b)$ on which f has finite variation.

Suppose instead that S is a closed set having positive measure. Exercise 797, which follows the proof, shows exactly how to choose a null subset Z of S that is a $\mathcal{G}_\delta$ -set that is dense in S . By our assumption (c), there must be a portion $Z \cap (a, b)$ on which f has σ -finite variation. We apply Theorem 9.20 to obtain a sequence of compact sets $\{K_n\}$ whose union includes $Z \cap (a, b)$ so that each $\lambda_f(K_n) < \infty$.

Apply the Osgood-Baire theorem, now to the sequence of compact sets $\{K_n\}$ that covers the $\mathcal{G}_\delta$ -set $Z \cap (a, b)$. Recall that the Osgood-Baire theorem, stated in Section 11.1.6 for closed sets, applies equally well to $\mathcal{G}_\delta$ -sets. Thus we can conclude that there is a portion $Z \cap (c, d)$ and an integer k so that $Z \cap (c, d) \subset K_k$. Since Z is dense in the compact set S we also have $S \cap (c, d) \subset K_k$. In particular

$$\lambda_f(S \cap (c, d)) \leq \lambda_f(K_k) < \infty.$$

We have obtained again (but this time without the additional assumption that S has measure zero) exactly property (b). ■

Exercise 797 Let S be a compact set. Show that there is a subset Z of S that is of type $\mathcal{G}_\delta$, is a null set, and is dense in S . Answer □

9.3.2 λ -absolutely continuous functions

As a corollary to Theorem 9.21 immediately we have a special observation, since an λ -absolutely continuous function must have finite variation (indeed zero variation) on every set of λ -measure zero.

Corollary 9.22 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function that is λ -absolutely continuous on a compact set E . Then f has σ -finite variation there.

9.4 Vitali property and differentiability

In this section we show that differentiability on a set implies the Vitali property on that set and, conversely, that the Vitali property on a set implies almost everywhere differentiability.

Theorem 9.23 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ have a finite derivative at every point of a set E . Then f has the Vitali property on E and, moreover,

$$\lambda_f(E) = \lambda_f^*(E) = \int_E |f'(x)| dx.$$

Proof. This is already proved in Theorem 9.8. ■

Theorem 9.24 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function that has the Vitali property on a set E . Then f has a finite derivative at almost every point of E and, except at the points of a set N for which $\lambda_f(N) = 0$, f has a finite or infinite derivative $f'(z)$.

Proof. We need work only with the Lipschitz numbers here. Recall that if $\underline{\text{lip}}_f(z) = \infty$ then necessarily f has an infinite derivative, $f'(z) = \infty$ or $f'(z) = -\infty$ (see Exercise 790). Also if

$$\underline{\text{lip}}_f(z) = \overline{\text{lip}}_f(z) < \infty$$

then f has a finite derivative at z (see Exercise 789).

It is enough to prove the theorem under the assumption that E is a bounded set. We examine

$$A = \{x \in E : \underline{\text{lip}}_f(x) < \overline{\text{lip}}_f(x)\}.$$

As is usual in arguments of this type, introduce rational numbers $0 < r < s$ and the subsets

$$A_{rs} = \{x \in A : \underline{\text{lip}}_f(x) < r < s < \overline{\text{lip}}_f(x)\}.$$

Note that A is the countable union of this collection of sets taken over all rationals r and s with $r < s$.

By the growth lemmas of Section 9.2.4 we obtain

$$\lambda_f^*(A_{rs}) \leq r\lambda(A_{rs}) \leq s\lambda(A_{rs}) \leq \lambda_f(A_{rs}).$$

Our assumption that f has the Vitali property on E gives the identity $\lambda_f = \lambda_f^*$ on each of these subsets of E . None of these numbers are infinite, $r < s$, and so the inequality makes sense only in the case that $\lambda_f(A_{rs}) = \lambda(A_{rs}) = 0$. Consequently $\lambda_f(A) = \lambda(A) = 0$.

At every point x in $E \setminus A$ we know that either

$$\underline{\text{lip}}_f(x) = \overline{\text{lip}}_f(x) < \infty$$

or else

$$\underline{\text{lip}}_f(x) = \overline{\text{lip}}_f(x) = +\infty.$$

In the former case, as we have already noted, f has a finite derivative and in the latter case f has an infinite derivative. This latter case can occur only on a set of Lebesgue measure zero (as a consequence of Lemma 9.19). ■

Corollary 9.25 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function that has the Vitali property on a set E and let us specify the following subsets of E at which the derivative exists finitely or infinitely:*

1. $E_d = \{x \in E : f \text{ is differentiable at } x\}.$
2. $E_\infty = \{x \in E : f'(x) = \pm\infty\}.$

Then

$$\lambda_f^*(E) = \lambda_f(E) = \int_{E_d} |f'(x)| dx + \lambda_f(E_\infty).$$

9.5 The Vitali property and variation

The Vitali property is closely related to the finiteness of the variation. Indeed, since the fine variation λ_f^* of a continuous function f is always σ -finite, we know that the identity $\lambda_f^*(E) = \lambda_f(E)$ can only hold if f has σ -finite variation on E .

9.5.1 Monotonic functions

Theorem 9.26 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous, strictly increasing function. Then f has the Vitali property.*

Proof. Theorem 9.6 supplies the identity

$$\lambda_f^*(E) = \lambda_f(E) = \lambda(f(E)).$$

■

Theorem 9.27 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous, monotonic nondecreasing function. Then f has the Vitali property.*

Proof. Let $\varepsilon > 0$ and define a new function $g(x) = f(x) + \varepsilon x$. The function g is continuous and strictly increasing so, by the previous theorem, $\lambda_{g*} = \lambda_g^*$. From Exercise 760 we deduce the inequalities

$$\lambda_f \leq \lambda_g^* \leq \lambda_f + \varepsilon \lambda^*$$

and

$$\lambda_f^* \leq \lambda_{g*} \leq \lambda_f^* + \varepsilon \lambda^*.$$

From these two inequalities and the identity $\lambda_{g*} = \lambda_g^*$ we can deduce $\lambda_f = \lambda_f^*$. ■

Exercise 798 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a monotonic, nondecreasing function. Show that if $\lambda_f^*(\{x\}) = \lambda_f(\{x\})$ for a point x then f must be continuous at x .* □

9.5.2 Functions of bounded variation

Theorem 9.28 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function that is locally of bounded variation. Then f has the Vitali property on the real line.*

Proof. Fix a compact interval $[a, b]$ and let g be the total variation function of f on $[a, b]$. We know that this relation between a function and its total variation function requires the identity

$$V^*(\Delta g - |\Delta f|, (a, b)) = 0.$$

In particular $\lambda_f(E) = \lambda_g(E)$ and $\lambda_f^*(E) = \lambda_g^*(E)$ for all subsets E of (a, b) . By the previous theorem $\lambda_g(E) = \lambda_g^*(E)$ and so $\lambda_f(E) = \lambda_f^*(E)$ follows. This argument produces the identity we require on all bounded sets, and the extension to arbitrary sets follows from measure properties. ■

9.5.3 Functions of σ -finite variation

Theorem 9.29 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function. Then f has σ -finite variation on a set E if and only if f has the Vitali property on E .*

Proof. We already know that the Vitali property for a continuous function will imply σ -finite variation. Let us prove the converse.

Suppose that f is continuous function that has σ -finite variation on E . By Theorem 9.20 there is a sequence of compact sets $\{E_n\}$ covering E and a sequence of functions g_n each continuous and locally of bounded variation so that

$$V^*(\Delta f - \Delta g_n, E_n) = 0 \quad (9.3)$$

We know then, from the previous theorem, that $\lambda_{g_n*} = \lambda_{g_n}^*$. We also know that the equivalence (9.3) requires that $\lambda_{g_n}^* = \lambda_f$ and $\lambda_{g_n*} = \lambda_f^*$ on all subsets of E_n .

Introduce the notation

$$A_n = E_n \setminus \bigcup_{k < n} E_k$$

so that $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} E_n$ and the sets $\{A_n\}$ are pairwise disjoint, measurable sets. The student should justify that the following computations are permitted:

$$\begin{aligned} \lambda_f(E) &= \sum_{n=1}^{\infty} \lambda_f(E \cap A_n) = \sum_{n=1}^{\infty} \lambda_{g_n}^*(E \cap A_n) = \\ &= \sum_{n=1}^{\infty} \lambda_{g_n*}(E \cap A_n) = \sum_{n=1}^{\infty} \lambda_f^*(E \cap A_n) = \lambda_f^*(E). \end{aligned}$$

As this applies as well to any subset of E we see that f must have the Vitali property on E as required. ■

Corollary 9.30 *If $f : \mathbb{R} \rightarrow \mathbb{R}$ is a continuous function that is λ -absolutely continuous on a compact set E , then f has the Vitali property on E .*

Proof. Use Corollary 9.22. ■

9.6 Characterization of the Vitali property

The class of functions satisfying the Vitali property on a set is fundamental to an understanding of the calculus program demanding the relation among the concepts of derivative, integral and variation. We have already found a number of characterizations in Theorem 9.20 and Theorem 9.21. Here are some more. Some are easy consequences of what we have proved [e.g., (a) and Theorem 775 immediately imply (b)]. Others are left as entertainments for the student.

Theorem 9.31 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous real function and let E be a compact set. The following are equivalent:*

1. f has the Vitali property on E .
2. f has σ -finite variation on E .
3. there is a sequence of compact sets $\{E_n\}$ with $E = \bigcup_{n=1}^{\infty} E_n$ so that for each n there is a continuous function g_n that is locally of bounded variation so that f and g_n are Kolmogorov equivalent on E_n .
4. f has a derivative (finite or infinite) at λ_f -almost every point of E .
5. There is a continuous, increasing function g so that

$$\limsup_{(I,x) \Rightarrow x} \left| \frac{\Delta f(I)}{\Delta g(I)} \right| < \infty$$

at every point $x \in E$.

6. There is a continuous, increasing function g and a real function f_1 so that

$$V^*(\Delta f - f_1 \Delta g, E) = 0.$$

7. There is a continuous, increasing function g so that the composed function $f \circ g$ has a finite derivative everywhere in the compact set $g^{-1}(E)$.

9.7 Characterization of λ -absolute continuity

The Vitali property expresses the most important property arising in studies of the derivative in the calculus. The special subclass of λ -absolutely continuous functions plays its most significant role in the integration theory. Here are some similar characterizations for this class, most easily proved from previously proved statements or techniques.

Theorem 9.32 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function and let E be a compact set. The following are equivalent:

1. f is λ -absolutely continuous on E .
2. f has σ -finite variation on E and is λ -absolutely continuous there.
3. there is a sequence of compact sets $\{E_n\}$ with $E = \bigcup_{n=1}^{\infty} E_n$ so that for each n there is a continuous function g_n that is of locally of bounded variation and absolutely continuous in the sense of Vitali so that f and g_n are Kolmogorov equivalent on E_n .
4. f has a finite derivative at λ_f -almost every point of E .
5. There is an increasing, λ -absolutely continuous function g so that

$$\limsup_{(I,x) \rightarrow x} \left| \frac{\Delta f(I)}{\Delta g(I)} \right| < \infty$$

at every point $x \in E$.

6. There is an increasing, λ -absolutely continuous function g and a real function f_1 so that

$$V^*(\Delta f - f_1 \Delta g, E) = 0.$$

9.8 Mapping properties

For any set E and any function $f : \mathbb{R} \rightarrow \mathbb{R}$ the image of E under the mapping f is written as

$$f(E) = \{f(x) : x \in E\}.$$

We already know some properties of the image set for continuous functions. We recall from elementary studies (e.g., Chapter 1) that the image of any compact interval $[a, b]$ under f is again a compact interval. It is easy to check that the image of any compact set E under f is again a compact set $f(E)$. A natural question is whether the image of an measurable set must also be measurable .

Theorem 9.33 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be an measurable function and P an measurable set. The following are equivalent:

(M) $f(E)$ is measurable for every measurable subset E of P ,

(N) $\lambda(f(N)) = 0$ for every subset N of P for which $\lambda(N) = 0$.

Proof. Suppose that E is measurable and that the second statement of the theorem holds. We need consider only the case where E is bounded. Since f is measurable, then by definition, we can find open sets G_n so that $\lambda(G_n) < 1/n$, $E \setminus G_n$ is compact and f is equal to a continuous function $g_n : \mathbb{R} \rightarrow \mathbb{R}$ on the compact set $E \setminus G_n$.

In particular

$$E = Z \cup \bigcup_{n=1}^{\infty} (E \setminus G_n)$$

where

$$Z = E \cap \bigcap_{n=1}^{\infty} G_n$$

has λ -measure zero. By hypothesis $f(Z)$ must be a set of λ -measure zero and hence is measurable. Also each

$$f(E \setminus G_n) = g_n(E \setminus G_n)$$

is a compact set (since the continuous function g_n maps compact sets to compact sets). In particular each set here is also measurable. Thus

$$f(E) = f(Z) \cup \bigcup_{n=1}^{\infty} f(E \setminus G_n)$$

displays $f(E)$ as the union of a sequence of measurable sets. Thus $f(E)$ is also measurable.

Conversely suppose that the first statement of the theorem does not hold, yet the second does. Then there is a set $Z \subset P$ for which $\lambda(Z) = 0$ and yet $f(Z)$ does not have λ -measure zero. For (b) to be true, however, $f(Z)$ should be an measurable set of positive measure. Such a set must have a subset A that is not measurable.

We shall not pause to prove this assertion but leave it as a project for the student to find elsewhere (or prove). A proof will require use of a logical principle that is beyond our elementary calculus course.

Then there is a set $Z_1 \subset Z$ with $f(Z_1) = A$. The set Z_1 must be measurable merely because $\lambda(Z_1) \leq \lambda(Z) = 0$. But

then f maps an measurable set Z_1 to a set $f(Z_1) = A$ that is not measurable. We have contradicted the second statement thus completing the proof. ■

9.9 Lusin's conditions

Definition 9.34 A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is said to satisfy Lusin's conditions on a set P when these equivalent conditions hold:

- (M) $f(E)$ is measurable for every measurable subset E of P ,
- (N) $\lambda(f(N)) = 0$ for every subset N of P for which $\lambda(N) = 0$.

Theorem 9.35 If $f : \mathbb{R} \rightarrow \mathbb{R}$ is λ -absolutely continuous on an measurable set P then f satisfies Lusin's conditions on P .

Proof. This follows immediately from Theorem 9.7 that asserts that $\lambda(f(N))$ is smaller than the full variation of f on N . Thus for every null set $N \subset P$,

$$\lambda(f(N)) \leq \lambda_f(N) = 0.$$

■

9.10 Banach-Zarecki Theorem

In the converse direction we should expect that Lusin's conditions play a role in characterizing the important property of absolute continuity.

Theorem 9.36 (Banach-Zarecki) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function and E a compact set. Then the following are necessary and sufficient conditions in order that f is λ -absolutely continuous on E :

1. f has σ -finite variation on E , and
2. f satisfies Lusin's condition on E .

Proof. Certainly if f is λ -absolutely continuous then we already know that (a) holds because of Theorem 9.21 and that (b) holds because of Theorem 9.35.

Conversely let us suppose that (a) and (b) now hold. We know from Theorem 9.20 that when f has σ -finite variation on a compact set E , there is a sequence $\{E_n\}$ of compact sets covering E and a sequence of continuous functions of bounded variation g_n so that f and g_n are Kolmogorov equivalent on E_n . Recall in the proof that the construction there required $f = g_n$ on the set E_n . We can insist on that here. Moreover the functions g_n in the proof that extended f were also chosen to be merely linear or constant in the intervals complementary to E_n . We can insist also on that here.

We note that the condition (b) of the theorem asserting that f satisfies Lusin's condition on E means that g_n satisfies this same condition on E_n . Moreover by the nature of the construction the function g_n satisfies Lusin's condition on all sets. The proof is completed now by addressing the special case of proving that g_n is λ -absolutely continuous.

Note that each g_n constructed in our proof above satisfies the hypotheses of Exercise 799 below. Indeed, since g_n has bounded variation on every interval it is differentiable outside of a set N of λ -measure zero. The assumption of Lusin's condition on g_n then provides $\lambda(g_n(N)) = 0$. The finiteness of

$$\lambda_{g_n}(\mathbb{R} \setminus N) = \int_{\mathbb{R} \setminus N} |g'_n(x)| dx$$

follows from the fact that g_n , as constructed have finite variation.

Now let Z be any set for which $\lambda(Z) = 0$. Let $\varepsilon > 0$ and choose $\delta > 0$ by applying the Exercise 799 to this function g_n . Choose an open set $G \subset Z$ with $\lambda(G) < \delta$. Choose any full cover β of Z ; then $\beta(G)$ is also a full cover of Z and the exercise provides

$$V^*(\Delta g_n, Z) \leq V(\Delta g_n, \beta(G)) < \varepsilon.$$

From this we deduce that $\lambda_{g_n}(Z) = 0$. In consequence g_n is λ -absolutely continuous.

From this we can prove that f is λ -absolutely continuous on the set E in question. For if Z is a set of λ -measure zero then $\lambda^{g_n}(Z) = 0$ will imply that

$$\lambda_f(E_n \cap Z) = \lambda^{g_n}(E_n \cap Z) = 0$$

and hence that

$$\lambda_f(E \cap Z) \leq \sum_{n=1}^{\infty} \lambda_f(E_n \cap Z) = 0.$$

This will then show that f is λ -absolutely continuous on E . ■

Corollary 9.37 Let $f : [a, b] \rightarrow \mathbb{R}$. The following are necessary and sufficient conditions in order that f is absolutely continuous in the sense of Vitali on $[a, b]$:

1. f is continuous,
2. f has bounded variation on $[a, b]$, and
3. f satisfies Lusin's conditions on $[a, b]$.

A crucial step in the proof of the theorem uses the following classical problem:

Exercise 799 Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function. Suppose that g is differentiable at each point with the exception of points in a set N for which $\lambda(g(N)) = 0$ and suppose that $\int_{\mathbb{R} \setminus N} |g'(x)| dx < \infty$. Show that, for every $\varepsilon > 0$, there is a $\delta > 0$ so that any sequence of nonoverlapping intervals $\{[c_n, d_n]\}$ for which $\sum_n \lambda([c_n, d_n]) < \delta$ it follows that

$$\sum_n |\Delta g([c_n, d_n])| < \varepsilon.$$

Answer $\square$

9.11 Local Lebesgue integrability conditions

A measurable function f is Lebesgue integrable on an interval $[a, b]$ provided that the integral $\int_a^b |f(x)| dx$ is finite. If the integral is not finite then f cannot be Lebesgue integrable on $[a, b]$. But need it be Lebesgue integrable on some subinterval? The theorem we now prove gives a sufficient condition in order for a measurable functions to have a local integrability property. In the theorem we use the following notation for a function f and a closed set E : the function f_E is defined as $f_E(x) = f(x)$ whenever $x \in E$ and $f_E(x) = 0$ otherwise.

Theorem 9.38 Let E be a nonempty closed subset of $[a, b]$ and f an measurable function. Suppose that

$$-\infty < \int_a^b f(x) dx \leq \int_a^b f(x) dx < \infty.$$

Then E contains a portion $E \cap (c, d)$ so that f_E is Lebesgue integrable on $[c, d]$.

Proof. We make a simplifying assumption that allows a small technical detail later. We remove from the set E all points that are isolated on either the right side or the left side or both sides. There are only countably many such points and that does not influence either measure or integration statements. While the resulting set is not closed, it is a set of type $\mathcal{G}_\delta$ so that we may still apply the Baire-Osgood theorem to it.

Choose t so that

$$-t < \int_a^b f(x) dx \leq \overline{\int_a^b f(x) dx} < t$$

and a Cousin cover β of $[a, b]$ so that³

$$\left| \sum_{\pi} f\lambda \right| < t$$

for all partitions $\pi \subset \beta$ of $[a, b]$. Let $[c, d]$ be any subinterval and let $\pi \subset \beta$ be a partition of $[c, d]$. Choose $\pi' \subset \beta$ so that it consists of a partition of $[a, c]$ and $[d, b]$. Then

$$\left| \sum_{\pi \cup \pi'} f\lambda \right| < t$$

so that

$$\left| \sum_{\pi} f\lambda \right| \leq t + \left| \sum_{\pi'} f\lambda \right|$$

In particular we can write

$$T(c, d) = \sup \left\{ \left| \sum_{\pi} f\lambda \right| : \pi \subset \beta \text{ is a partition of } [c, d] \right\} < \infty.$$

We need a decomposition argument for β similar to that in Section 5.1.8. Choose $\delta(x) > 0$ so that $x \in I \subset [a, b]$ and $\lambda(I) < 2\delta(x)$ requires $(I, x) \in \beta$. Define

$$E_n^+ = \{x \in E : \delta(x) > 1/n, 0 \leq f(x) \leq n\}$$

³We are using $\sum_{\pi} f\lambda$ to denote the sum $\sum_{(I,x) \in \pi} f(x)\lambda(I)$ in this proof as many such sums will be considered.

and

$$E_n^- = \{x \in E : \delta(x) > 1/n, 0 \geq f(x) \geq -n\}.$$

This sequence of sets exhausts the set E so that, by the Baire-Osgood theorem, there must be a portion of E so that one of the sets is dense there. Thus we are able to choose an integer m and a subinterval $[c, d]$ so that $d - c < 1/m$ and so that E_m^+ (say) is dense in the nonempty portion $E \cap (c, d)$.

We shall investigate the Lebesgue integrability of f_E on $[c, d]$. For that, let π be an arbitrary partition of $[c, d]$ chosen from β . We shall estimate

$$\sum_{\pi} f_E^+ \lambda \quad \text{and} \quad \sum_{\pi} f_E^- \lambda$$

(where, as usual, f_E^+ and f_E^- denote the positive and negative parts of f_E).

Define $\pi_1 = \pi[E]$ and $\pi_2 = \pi \setminus \pi_1$. We alter π_1 in two different ways. The first alteration denoted as π_1' will replace each $(I, x) \in \pi_1$ by (I, x') where $x' \in E_m^+$. Since $x \in E$ and is not isolated on either side in E , and since E_m^+ is dense in this portion of E , such points are available. For any such point x' we see that the pair $(I, x') \in \beta$ because $\lambda(I) < 1/m < \delta(x')$. The second alteration denoted as π_1'' will replace each $(I, x) \in \pi_1$ for which $f(x) < 0$ by (I, x'') where $x'' \in E_m^+$. For the same reasons as before, the pair $(I, x'') \in \beta$. We will make use of the fact that, for the adjusted points x' and x'' , we have the inequalities $0 \leq f(x') \leq m$ and $f(x) < 0 \leq f(x'')$.

Now we do our computations:

$$\left| \sum_{\pi_1 \cup \pi_2} f \lambda \right| \leq T(c, d) \tag{9.4}$$

$$\left| \sum_{\pi_1' \cup \pi_2} f \lambda \right| \leq T(c, d) \tag{9.5}$$

$$\left| \sum_{\pi_1'} f \lambda \right| \leq m(d - c) \leq 1. \tag{9.6}$$

Combining (9.5) and (9.6) we see that

$$\left| \sum_{\pi_2} f \lambda \right| \leq T(c, d) + 1 \tag{9.7}$$

Thus we can estimate

$$\begin{aligned} \sum_{\pi} f_E^+ \lambda &= \sum_{\pi_1} f_E^+ \lambda = \sum_{\pi_1''} f_E^+ \lambda \leq \sum_{\pi_1''} f \lambda \\ &\leq \sum_{\pi_1''} f \lambda + \left[\sum_{\pi_2} f \lambda + T(c, d) + 1 \right] = \sum_{\pi_1'' \cup \pi_2} f \lambda + T(c, d) + 1 \leq 2T(c, d) + 1. \end{aligned}$$

As such sums have this upper bound we can conclude that

$$\overline{\int_c^d f_E^+(x) dx}$$

is finite and hence that the measurable function f_E^+ is Lebesgue integrable on $[c, d]$.

Now we show that f_E^- is also Lebesgue integrable on $[c, d]$. Since

$$f_E^-(x) = f_E^+(x) - f(x)$$

for every $x \in E$, we find that

$$\begin{aligned} \sum_{\pi} f_E^- \lambda &= \sum_{\pi_1} f_E^- \lambda = \sum_{\pi_1} f_E^+ \lambda - \sum_{\pi_1} f \lambda \\ &= \left[\sum_{\pi_1} f_E^+ \lambda + \sum_{\pi_2} f_E^+ \lambda \right] - \sum_{\pi_1} f \lambda \\ &\leq [2T(c, d) + 1] - \sum_{\pi_1} f \lambda - \left[\sum_{\pi_2} f \lambda - T(c, d) - 1 \right] \\ &= [3T(c, d) + 2] - \sum_{\pi} f \lambda \leq 4T(c, d) + 2. \end{aligned}$$

Once again such sums have this upper bound we can conclude that the measurable function f_E^- is Lebesgue integrable on $[c, d]$. Finally then $f_E = f_E^+ + f_E^-$ too must be Lebesgue integrable on $[c, d]$. This gives us our portion $E \cap (c, d)$ and completes the proof. ■

9.12 Continuity of upper and lower integrals

The indefinite integral of an integrable function is continuous. We can express this by saying that, if f is integrable on a compact interval $[a, b]$, then for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$-\varepsilon < \int_c^d f(x) dx < \varepsilon$$

for every subinterval $[c, d] \subset [a, b]$ for which $\lambda([c, d]) < \delta$. We wish a version of this that does not assume integrability and that can be used for a characterization.

Definition 9.39 A function f is said to have continuous upper and lower integrals on a compact interval $[a, b]$ if for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$-\varepsilon < \int_c^d f(x) dx \leq \overline{\int_c^d f(x) dx} < \varepsilon$$

for every subinterval $[c, d] \subset [a, b]$ for which $\lambda([c, d]) < \delta$.

Lemma 9.40 Suppose that $f : [a, b] \rightarrow \mathbb{R}$ has continuous upper and lower integrals on a compact interval $[a, b]$. Then

$$-\infty < \int_c^d f(x) dx \leq \overline{\int_c^d f(x) dx} < \infty$$

for every subinterval $[c, d] \subset [a, b]$.

Proof. There must be a $\delta > 0$ so that

$$-1 < \int_c^d f(x) dx \leq \overline{\int_c^d f(x) dx} < 1$$

for every subinterval $[c, d] \subset [a, b]$ for which $\lambda([c, d]) < \delta$. Subdivide

$$a = a_0 < a_1 < \cdots < a_{n-1} < a_n = b$$

9.13. A CHARACTERIZATION OF THE INTEGRAL

387

in such a way that each $a_i - a_{i-1} < \delta$. Then compute, using Exercise 647, that

$$\int_a^b f(x) dx = \sum_{i=1}^n \int_{a_{i-1}}^{a_i} f(x) dx \leq n < \infty.$$

A similar argument handles the lower integral. ■

Exercises

Exercise 800 (Cauchy extension property) Let f be integrable on every subinterval $[c, d] \subset (a, b)$. Show that f is integrable on $[a, b]$ if and only if f has continuous upper and lower integrals on $[a, b]$. Answer □

Exercise 801 (Harnack extension property) Let $F : \mathbb{R} \rightarrow \mathbb{R}$, let E be a closed subset of $[a, b]$, and let $\{(a_i, b_i)\}$ be the sequence of intervals complementary to E in (a, b) . Suppose that

1. $f(x) = 0$ for all $x \in E$,
2. f is integrable on all intervals $[a_i, b_i]$, and
3. $\sum_{i=1}^{\infty} \sup_{a_i \leq c_i < d_i \leq b_i} \left| \int_{c_i}^{d_i} f(x) dx \right| < \infty$.

Show that f is integrable on $[a, b]$ and

$$\int_a^b f(x) dx = \sum_{i=1}^{\infty} \int_{a_i}^{b_i} f(x) dx.$$

Answer □

9.13 A characterization of the integral

The class of Lebesgue integrable functions on an interval $[a, b]$ can be characterized as those measurable functions f for which

$$\int_a^b |f(x)| dx < \infty.$$

We now show that the full class of integrable functions (absolutely or nonabsolutely) on an interval $[a, b]$ can be characterized as those measurable functions that have continuous upper and lower integrals.

Theorem 9.41 *A function f is integrable on $[a, b]$ if and only if f is measurable and f has continuous upper and lower integrals on $[a, b]$.*

Proof. We already know that an integrable function has these properties. Conversely suppose that f is measurable and that f has continuous upper and lower integrals on $[a, b]$. An open interval $(s, t) \subset (a, b)$ will be called “accepted” if f is integrable on every $[c, d] \subset (s, t)$. Let G be the union of all accepted intervals. This is an open subset of (a, b) . Note that, if $[c, d] \subset G$, then by the Heine-Borel property $[c, d]$ can be written as the union of a finite collection of intervals $\{[c_i, d_i]\}$ each of which is inside an accepted interval. It follows that f is integrable on $[c, d]$ too.

Let

$$G = \bigcup_{i=1}^{\infty} (a_i, b_i),$$

displaying G as a union of its component intervals. We claim first that f must be integrable on each of the compact intervals $[a_i, b_i]$. This follows directly from the Cauchy extension property (Exercise 800) using the hypothesis that f has continuous upper and lower integrals. We shall use a single function F to represent the indefinite integral of f on each of these intervals, but we are cautioned not to use F outside of the intervals.

In particular if $G = (a, b)$ then the proof is completed since then f must be integrable on $[a, b]$ as required. Suppose not, i.e., that the theorem fails and $G \neq (a, b)$. Then $E = [a, b] \setminus G$ is a nonempty closed set. Note that E can have no isolated points. Indeed if $c \in E$ is isolated then $(c - t, c) \subset G$ and $(c, c + t) \subset G$ for some $t > 0$ and another application of the Cauchy extension property would show that $(c - t, c + t)$ is accepted so that $(c - t, c + t) \subset G$ which is not possible.

The goal of the proof now will be to obtain a portion $E \cap (c', d')$ of E with the property that (c', d') is accepted, which would be impossible. Portions cannot be empty and no point of E would be allowed to belong to an accepted interval. The local integrability Theorem 9.38 and the Harnack extension property (Exercise 801) will play key roles.

The assumption that f satisfies the continuity condition in Definition 9.39 together with Lemma 9.40 shows that the upper and lower integrals of f are finite. Thus, we can apply Theorem 9.38 to find a portion $E \cap [c, d]$ so that f_E is Lebesgue integrable on $[c, d]$.

Since f has continuous upper and lower integrals on $[c, d]$ it follows from Lemma 9.40 that

$$-\infty < \int_c^d f(x) dx \leq \overline{\int_c^d f(x) dx} < \infty.$$

Since f_E is Lebesgue integrable on $[c, d]$ it follows that

$$\int_c^d |f_E(x)| dx < \infty.$$

Thus we can select a real number $M > 0$ and a Cousin cover β of $[c, d]$ so that for any partition π of $[c, d]$ from β both

$$\left| \sum_{\pi} f \lambda \right| < M$$

and

$$\sum_{\pi} |f_E| \lambda < M.$$

We need a decomposition argument for β similar to that in Section 5.1.8. Choose $\delta(x) > 0$ so that $x \in I \subset [a, b]$ and $\lambda(I) < 2\delta(x)$ requires $(I, x) \in \beta$. Define

$$E_n = \{x \in E \cap [c, d] : \delta(x) > 1/n\}.$$

This sequence of sets exhausts the set $E \cap [c, d]$ so that, by the Baire-Osgood theorem, there must be a portion of so that one of the sets is dense there. Let us agree that E_m is dense in $E \cap (c', d')$ and that $[c', d']$ is smaller in length than $1/m$. Let $\{(c_i, d_i)\}$ denote the component intervals of $(c', d') \setminus E$. There must be infinitely many such component intervals since otherwise it would follow that f is integrable on $[c', d']$. We claim that

$$\sum_{i=1}^{\infty} \omega F([c_i, d_i]) = \infty. \quad (9.8)$$

For, if not, then the Harnack extension property (Exercise 801) shows that $f - f_E$ must be integrable on $[c', d']$ and hence f is integrable there. But that contradicts the fact that $[c', d']$ must contain points of E .

From the continuity of F we know that

$$\omega F([c_i, d_i]) = |F(s) - F(t)| \quad (9.9)$$

for some subinterval $[s, t] \subset [c_i, d_i]$. Consequently we may choose a sequence of intervals $\{[s_k, t_k]\}$, chosen from different

component intervals $[c_i, d_i]$ in such a way that either

$$0 \leq \sum_{k=1}^{\infty} F(t_k) - F(s_k) = \infty \quad (9.10)$$

or

$$0 \geq \sum_{k=1}^{\infty} F(t_k) - F(s_k) = -\infty. \quad (9.11)$$

Let us assume the former. If (9.11) holds instead the same argument with a slight adjustment in the inequalities will work.

Now we fix an integer p and carefully construct a partition π of the interval $[c, d]$ from β . The first step is to choose π' from β to form a partition of $[c, c']$, then π'' from β to form a partition of $[d', d]$. For each of the intervals $\{[s_k, t_k]\}$ for $k = 1, 2, 3, \dots, p$ we select a partition π_k of $[s_k, t_k]$ in such a way that

$$|F(t_k) - F(s_k) - \sum_{\pi_k} f\lambda| < 2^{-k}. \quad (9.12)$$

This is possible since f is integrable on each such interval and F is an indefinite integral. To complete the partition we take the remaining intervals, not yet covered by

$$\pi' \cup \pi'' \cup \bigcup_{k=1}^p \pi_k.$$

There are only finitely many of these intervals, say $I_1, I_2, \dots, I_q$. Each is a subinterval of $[c', d']$ and each one contains many points of E ; thus each one also contains a point of E_m . Select a point x_i from $E_m \cap I_i$ ($i = 1, 2, \dots, q$) and note that (I_i, x_i) belongs to β . Thus if we set

$$\pi''' = \{(I_i, x_i) : i = 1, 2, \dots, q\}$$

then we have obtained a partition

$$\pi = \pi' \cup \pi'' \cup \pi''' \cup \bigcup_{k=1}^p \pi_k$$

of the interval $[c, d]$ that is contained in β .

Consequently, by the way in which we chose M and β ,

$$\left| \sum_{\pi} f\lambda \right| \leq M.$$

We know too that

$$\left| \sum_{\pi'''} f\lambda \right| \leq \sum_{\pi'''} |f_E| \lambda \leq M.$$

We combine these inequalities with (9.12) and the simple inequality

$$\sum_{k=1}^p 2^{-k} \leq 1$$

to obtain

$$\sum_{k=1}^p F(t_k) - F(s_k) \leq \left| \sum_{\pi' \cup \pi''} f\lambda \right| + 2M + 1.$$

This is true for any p and conflicts with our assumption that the inequality (9.10) holds.

Since neither inequality (9.10) nor (9.11) can hold it follows that inequality (9.9) also fails, thus f is integrable on $[c', d']$. In other words (c', d') is accepted, which would be impossible. This completes the proof. ■

9.14 Integral of Dini derivatives

If F is a continuous function on an interval $[a, b]$ and has a finite Dini derivative, say $D_+F(x)$, at each point then f is determined up to an additive constant by that Dini derivative. One suspects that

$$F(x) - F(a) = \int_a^x D_+F(t) dt$$

but this is not necessarily true and even when it is true we need some further methods to handle.

9.14.1 Motivation

We require a variant on the Cousin covering lemma that is more appropriate for handling the Dini derivatives of continuous functions.

The Cousin covers are particularly suited to describing properties of the ordinary derivative. For example if $\underline{D}F(x) > r$ then the covering relation

$$\beta = \{(I, x) : \Delta F(I) > r\lambda(I)\}$$

has the property that for some $\delta > 0$, if $x \in I$ and $\lambda(I) < \delta$ then necessarily $(I, x) \in \beta$. Indeed $\underline{D}F(x) > r$ if and only if β has this property.

We conclude this chapter by determining how to recover a function from one of its Dini derivatives and so will require a one-sided analogue. The simplest version could come from the observation that $D_+F(x) > r$ if and only if the covering relation

$$\beta = \{(I, x) : \Delta F(I) > r\lambda(I)\}$$

has the property that for some $\delta > 0$, if $0 < h < \delta$ then necessarily $([x, x+h], x) \in \beta$.

But in fact our covering relation needs to be designed to handle the upper Dini derivative, not the lower. For that the description is more delicate: $D^+F(x) > r$ if and only if the covering relation

$$\beta = \{(I, x) : \Delta F(I) > r\lambda(I)\}$$

has the property that for any $\varepsilon > 0$, there is at least one value of h with $0 < h < \varepsilon$ for which $([x, x+h], x) \in \beta$. We strengthen this by insisting that F is continuous. In that case, if we found h so that

$$\frac{F(x+h) - F(x)}{(x+h) - x} > r,$$

notice that there must be a $\delta > 0$ so that

$$\frac{F(x'+h) - F(x)}{(x'+h) - x} > r$$

for every value of x' in the interval $[x - \delta, x]$.

Definition 9.42 Let K be a compact set with endpoints $a = \inf K$ and $b = \sup K$. A covering relation β is said to be a quasi-Cousin cover of K provided that

1. There is at least one pair $([a, d], a) \in \beta$ with $a < d \leq b$.
2. For every $a < x < b$, $x \in K$ there is a $\delta > 0$ so that there is at least one $x < d \leq b$ for which all pairs $([c, d], x) \in \beta$ whenever $x - \delta < c \leq x$.
3. There is a $\delta > 0$ so that all pairs $([c, b], b) \in \beta$ whenever $b - \delta < c < b$.

9.14.2 Quasi-Cousin covering lemma

Even though the notion of a quasi-Cousin cover is much weaker than that of a Cousin cover the covering lemma generalizes.

Lemma 9.43 (Quasi-Cousin covering lemma) Let β be a quasi-Cousin cover of a compact set K with endpoints $a = \inf K$ and $b = \sup K$. Then β contains a subpartition π so that

$$K \subset \bigcup_{(I,x) \in \pi} I \subset [a, b].$$

Proof. Let us assume first that $K = [a, b]$. Let E be the set of all points z , with $b \geq z > a$ and with the property that β contains a partition π of $[a, z]$.

Argue that (i) $E \neq \emptyset$, (ii) if $\sup E = t$ then t cannot be less than b , (iii) if $\sup E = b$ then $b \in E$.

We know that (i) is true since there is at least one pair $([a, d], a) \in \beta$ with $a < d \leq b$ and so $d \in E$. Thus we may set $t = \sup E$ and be assured that $a < d \leq t \leq b$. To see (ii) note that it is not possible for $t < b$ for if so then there is a $\delta > 0$ and $d' > t$ for which all pairs $(t, [c, d']) \in \beta$ with $t - \delta < c \leq t$. But that supplies a point $t' \in (t - \delta, t) \cap E$ and the partition of $[a, t']$ can be enlarged by including $(t, [t', d'])$ to form a partition of $[a, d']$; thus $d' \in E$. But this violates $t = \sup E$.

Finally for (iii) if $t = b$ and yet $b \notin E$ then, repeating much the same argument, there is a $\delta > 0$ for which all pairs $(b, [c, b]) \in \beta$ with $b - \delta < c < b$. But that supplies a point $t' \in (b - \delta, b) \cap E$ and the partition for $[a, t']$ can be enlarged by including $(b, [t', b])$ to form a partition π for $[a, b]$. This shows that $b \in E$ after all.

Now let us handle the general case for an arbitrary compact set $K \subset [a, b]$. Let $G = (a, b) \setminus K$ and

$$\beta_1 = \{(I, x) : x \in I \text{ and } I \subset G\}.$$

Since β is a quasi-Cousin cover of K we can check that $\beta \cup \beta_1$ is a quasi-Cousin cover of $[a, b]$. By the first part of the proof there is a partition $\pi \subset \beta \cup \beta_1$ of $[a, b]$. Remove those elements of π that do not belong to β to form a subpartition with exactly the required properties. ■

The proof contains explicitly the statement of the corollary:

Corollary 9.44 *Let β be a quasi-Cousin cover of a compact interval $[a, b]$. Then β contains a partition of $[a, b]$ (although not necessarily of other subintervals of $[a, b]$).*

Exercises

Exercise 802 (variant on the quasi-Cousin covering) *Let K be a compact set and β a covering relation. Suppose that, for each $x \in K$, there are $s, t > 0$ so that all pairs*

$$([x', x + s], x) \in \beta$$

whenever $x - t \leq x' \leq x$. Show that β contains a subpartition π for which

$$K \subset \bigcup_{(I, x) \in \pi} I.$$

□

Exercise 803 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be continuous at each point of an open interval (a, b) and suppose that $D^+ f(x) > m$ for each $x \in (a, b)$. Then $f(d) - f(c) > m(d - c)$ for each $[c, d] \subset (a, b)$.* □

9.14.3 Estimates of integrals from derivates

As a warm-up to our theorem about Dini derivatives let us show that the ordinary derivatives are easily handled.

Lemma 9.45 Let $F, f : [a, b] \rightarrow \mathbb{R}$. If F is continuous at a and b and

$$\underline{D}F(x) \geq f(x)$$

at every point of (a, b) , then

$$\int_a^b f(x) dx \leq F(b) - F(a).$$

Proof. Let $\varepsilon > 0$. Take the covering relation

$$\beta_1 = \{(I, x) : \Delta F(I) \geq (f(x) - \varepsilon)\lambda(I)\}$$

and

$$\beta_2 = \{(I, x) : x = a \text{ or } b, x \in I \text{ and } |\Delta F(I)| + |f(x)|\lambda(I) < \varepsilon\}.$$

Check that $\beta = \beta_1 \cup \beta_2$ is a Cousin cover of $[a, b]$. At the endpoints a or b the continuity of F needs to be used in the verification, while at the points in (a, b) the inequality $\underline{D}F(x) \geq f(x)$ is used.

Any partition $\pi \subset \beta$ of the interval $[a, b]$ will satisfy

$$\sum_{(I, x) \in \pi} f(x)\lambda(I) \leq \sum_{(I, x) \in \pi} [\Delta F(I) + \varepsilon\lambda(I)] + 2\varepsilon = F(b) - F(a) + \varepsilon(2 + b - a).$$

This is true for all partitions π from this β and all $\varepsilon > 0$ and so the conclusion that

$$\int_a^b f(x) dx \leq F(b) - F(a)$$

now follows. ■

Lemma 9.46 Let $F, f : [a, b] \rightarrow \mathbb{R}$. If F is continuous at a and b and

$$\overline{D}F(x) \leq f(x)$$

at every point of (a, b) , then

$$\int_a^b f(x) dx \geq F(b) - F(a).$$

Proof. Apply Lemma 9.45 to the functions $-F$ and $-f$. ■

9.14.4 Estimates of integrals from Dini derivatives

For Dini derivatives there is a weaker version of Theorem 9.45 available using similar arguments (but employing quasi-Cousin covers as well as Cousin covers). Note that this weaker version uses lower and upper rather than upper and lower integrals; in particular no corollary can be derived asserting the integrability of the Dini derivative (indeed it may not be integrable).

Theorem 9.47 Suppose that $F : [a, b] \rightarrow \mathbb{R}$ is continuous and that g is a finite-valued function. If $D^+F(x) \geq g(x)$ at every point $a < x < b$, then,

$$F(b) - F(a) \geq \int_a^b g(x) dx. \quad (9.13)$$

If $D_+F(x) \leq g(x)$ at every point $a < x < b$, then

$$F(b) - F(a) \leq \int_a^b g(x) dx. \quad (9.14)$$

Proof. Let $\varepsilon > 0$. Take the covering relation β_1 of all pairs $([x, y], z)$ with

$$\Delta F([x, y]) \geq (f(z) - \varepsilon)\lambda([x, y])$$

and β_2 of all pairs $([a, y], a)$ and $([x, b], b)$ for which

$$\Delta F([a, y]) - f(a)\lambda([a, y]) > -\varepsilon$$

and

$$\Delta F([x, b]) \geq f(b)\lambda([x, b]) - \varepsilon.$$

It is easy to verify that $\beta = \beta_1 \cup \beta_2$ is a quasi-Cousin cover of $[a, b]$. At the endpoints a or b the continuity of F needs to be used in the verification, while at the points in (a, b) the inequality $D^+F(x) \geq g(x)$ is used.

This may not seem too much of a help since the integral is defined by Cousin covers, not by quasi-Cousin covers. But let β_3 be any Cousin cover of $[a, b]$. Check that, as defined, $\beta_3 \cap \beta$ must be a quasi-Cousin cover of $[a, b]$. Thus there is at least one partition π from β_3 that is also in β . For that partition a familiar argument gives us

$$\sum_{(I, x) \in \pi} f(x)\lambda(I) \leq \sum_{(I, x) \in \pi} [\Delta F(I) + \varepsilon\lambda([x, y])] + 2\varepsilon = F(b) - F(a) + \varepsilon(2 + b - a).$$

Note that this means any Cousin cover of $[a, b]$ contains at least one partition π with this property. Thus, while we can say nothing about the *upper* integral, we certainly can assert that the *lower* integral must always be lesser than $F(b) - F(a) + \varepsilon(2 + b - a)$ and from this the theorem follows. ■

As a consequence of this theorem we observe that if an everywhere finite function g is *assumed* to be integrable on $[a, b]$ and lies between the two derivatives then an integral identity holds. The assumption that g is integrable cannot be dropped here.

Corollary 9.48 *Let $F : [a, b] \rightarrow \mathbb{R}$ be continuous and $g : [a, b] \rightarrow \mathbb{R}$ be integrable on $[a, b]$ and suppose that*

$$D_+F(x) \leq g(x) \leq D^+F(x)$$

at every point x on $[a, b]$. Then

$$F(b) - F(a) = \int_a^b g(x) dx.$$

Exercises

Exercise 804 *Suppose that $F : \mathbb{R} \rightarrow \mathbb{R}$ is a continuous function and that $D^+F(x) > r$ at every point of an interval $[a, b]$. Verify that the covering relation*

$$\beta = \{(I, x) : \Delta F(I) > r\lambda(I)\}$$

satisfies the first two conditions (but not necessarily the third) in Definition 9.42. □

Exercise 805 *Continuing the previous exercise, let $\varepsilon > 0$ and let*

$$\beta' = \{([x - t, x]), x) : |F(x - t) - F(x)| < \varepsilon\}.$$

Show that $\beta \cup \beta'$ is a quasi-Cousin cover of $[a, b]$. □

Exercise 806 *Show that every Cousin cover of an interval $[a, b]$ is also a quasi-Cousin cover for any compact subset of $[a, b]$.* □

Exercise 807 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and suppose that the function $D^+f(x)$ is finite-valued and continuous at a point x_0 . Show that f is differentiable at x_0 .* □

Chapter 10

Integration in $\mathbb{R}^n$

In this chapter we shall sketch a theory of integration for functions of several variables. This is just a sketch to illustrate that the methods developed in the text extend without too much trouble to higher dimensions. The reader is, by now, ready for a full treatment using any of the standard presentations but may find it convenient to see a rapid account extending some of our techniques here.

The exercises do the technical work and, for the most part we have been content to give references to where the techniques needed can be found. We consider this final chapter more of a guide to thinking about this subject and the exercises and discussions in the Answers section are more a dialogue than a course of study.

10.1 Some background

We must assume the reader is familiar with the rudiments of analysis in the space $\mathbb{R}^n$. In particular these facts will be used.

- $\mathbb{R}^n$ is the collection of all n -tuples of real numbers $x = (x_1, x_2, \dots, x_n)$.
- Addition in $\mathbb{R}^n$ is defined by

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n).$$

- Scalar multiplication in $\mathbb{R}^n$ is defined by

$$r(x_1, x_2, \dots, x_n) = (rx_1, rx_2, \dots, rx_n).$$

- Distances in $\mathbb{R}^n$ are defined by

$$\|(x_1, x_2, \dots, x_n) - (y_1, y_2, \dots, y_n)\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}.$$

- The open ball with center $x = (x_1, x_2, \dots, x_n)$ and radius r in $\mathbb{R}^n$ is

$$B(x; r) = \{y : \|x - y\| < r\}.$$

10.1.1 Intervals and covering relations

By a closed interval in $\mathbb{R}$ we mean, of course, the set

$$I = [a, b] = \{x : a \leq x \leq b\}.$$

That set has two endpoints and the interior is the open interval (a, b) between them. The symbol $|I|$ denotes the length of I , i.e., $|I| = b - a$.

By an interval in $\mathbb{R}^2$ we mean a product of two intervals in $\mathbb{R}$. Thus the closed rectangle

$$I = [a, b] \times [c, d] = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}.$$

That set has four vertices, (a, c) , (b, c) , (b, d) , and (a, d) . The symbol $|I|$ denotes the area of I , i.e., $|I| = (b - a)(d - c)$ which is the product of the length and width of the rectangle.

These ideas and notation extend without difficulty to any dimension greater than two. By an *interval* in $\mathbb{R}^n$ we shall mean a cartesian product of one-dimensional intervals. It will be a closed interval if it is a product of closed intervals. Thus

$$I = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]$$

is the set of points in $\mathbb{R}^n$ described by these inequalities:

$$\{(x_1, x_2, \dots, x_n) : a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2, \dots, a_n \leq x_n \leq b_n\}.$$

This interval has 2^n vertices. The symbol $|I|$ denotes the n -dimensional volume of I , i.e.,

$$|I| = (b_1 - a_1)(b_2 - a_2)(b_3 - a_3) \dots (b_n - a_n)$$

which is the product of the length of all the edges in the interval.

Two intervals are *nonoverlapping* if their intersection has no interior points. Thus nonoverlapping intervals are either disjoint or else they meet only at some boundary points. A *packing* is a finite covering relation

$$\{(I_1, x_1), (I_2, x_2), (I_3, x_3), \dots, (I_k, x_k)\}$$

where each I_i is an interval and x_i is a point in the corresponding interval I_i , and distinct pairs of intervals I_i and I_j do not overlap.

By a *full interval cover* of a set $E \subset \mathbb{R}^n$ we mean a covering relation β that consists of pairs (I, x) again for which each I is an interval and x is a point in the corresponding interval I , and which is full in the following (by now familiar) sense: for each $x \in E$ there is a positive $\delta(x)$ so that β contains every pair (I, x) for which I is an interval containing x and contained in the open ball $B(x; \delta(x))$.

Exercise 808 (additivity of the volume) Show that the n -dimensional volume is an additive interval function, i.e., show that if J is a closed interval in $\mathbb{R}^n$ and π a packing for which

$$J = \bigcup_{(I, x) \in \pi} I$$

then

$$|J| = \sum_{(I, x) \in \pi} |I|.$$

Answer ☐

Exercise 809 (Cousin's lemma) Show that if β is a full interval cover of a closed interval J in $\mathbb{R}^n$ then there is a packing $\pi \subset \beta$ for which

$$J = \bigcup_{(I, x) \in \pi} I.$$

Answer ☐

10.2 Measure and integral

The measure theory and the integration are defined by means of full interval covers and packings. This is the analogue of the Riemann sums expression that was available in dimension one for all of our integrals in the early chapters.

Definition 10.1 Let $E \subset \mathbb{R}^n$ and let f be a nonnegative real-valued function defined on E . Then we define the upper integral

$$\overline{\int_E} f(x) dx = \inf_{\beta} \sup_{\pi \subset \beta} \sum_{(I,x) \in \pi} f(x)|I|$$

where the supremum is with regard to all packings $\pi \subset \beta$ where β is an arbitrary full interval cover of E . We use also the notation

$$\mathcal{L}^n(E) = \overline{\int_E} dx$$

and refer to the set function $\mathcal{L}^n$ as Lebesgue measure in $\mathbb{R}^n$.

The reader might well have expected a higher dimensional integral to look more like the one-dimensional version. For example if $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ perhaps we would expect an indefinite integral $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ defined as

$$F(x,y) = \int_a^x \int_b^y f(s,t) ds dt.$$

But the theory is far better expressed by the set function

$$E \rightarrow \overline{\int_E} f(x) dx$$

and it is this idea and notation that we pursue.

Note that if E is a bounded set then the upper integral could have been simply stated as an interval function by noticing that

$$\overline{\int_I} f(x) \chi_E(x) dx = \overline{\int_E} f(x) dx$$

for every interval I that contains E . Thus the theory could have been developed by Riemann sums over partitions of intervals. We prefer to pass immediately to the set version $E \rightarrow \overline{\int_E} f(x) dx$ which is closer to the mainstream of integration theory.

We shall not introduce a lower integral (as might be expected) but we will instead define what is meant by a $\mathcal{L}^n$ -measurable set and a $\mathcal{L}^n$ -measurable function. When E is a $\mathcal{L}^n$ -measurable set and a f is a $\mathcal{L}^n$ -measurable function then the Lebesgue integral

$$\int_E f(x) dx$$

will be defined to be the value

$$\overline{\int_E [f(x)]^+ dx} - \overline{\int_E [f(x)]^- dx}$$

provided this has a meaning (i.e., is not $\infty - \infty$). Thus the upper integral will serve us only as a tool that leads quickly to a formal expression for the value of the Lebesgue integral and the Lebesgue measure.

10.2.1 Lebesgue measure in $\mathbb{R}^n$

We use the special notion

$$\mathcal{L}^n(E) = \overline{\int_E dx}$$

and refer to this as n -dimensional *Lebesgue [outer] measure* on $\mathbb{R}^n$. This is defined for all subsets E of $\mathbb{R}^n$ as is the upper integral

$$\overline{\int_E f(x) dx}$$

which is defined for all functions f that assign a nonnegative number at every point of the set E .

We shall discover that for intervals $\mathcal{L}^n(I) = |I|$ so that Lebesgue measure is an extension of the volume function from the class of closed intervals to the class of all subsets of $\mathbb{R}^n$. Some authors prefer to keep the same notation in which case $|E|$ is defined for all subsets of $\mathbb{R}^n$ as

$$|E| = \overline{\int_E dx}.$$

10.2.2 The fundamental lemma

The fundamental lemma that we need that describes the key property of the upper integral and the measure is the following, seen already in its one-dimensional version in Lemma 6.26. The same proof works here to give essentially

the same conclusion.

Lemma 10.2 *Let $E \subset \mathbb{R}^n$ and let $f, f_1, f_2, f_3, \dots$ be a sequence of nonnegative real-valued functions defined on E . Suppose that*

$$f(x) \leq \sum_{k=1}^{\infty} f_k(x)$$

for every $x \in E$. Then

$$\overline{\int_E f(x) dx} \leq \sum_{k=1}^{\infty} \overline{\int_E f_k(x) dx}.$$

The two corollaries follow immediately and show that the set functions

$$E \rightarrow \overline{\int_E} dx$$

and

$$E \rightarrow \overline{\int_E} f(x) dx$$

are measures on $\mathbb{R}^n$ in the sense we make precise in Section 10.4 below.

Corollary 10.3 *Let $E, E_1, E_2, E_3, \dots$ be a sequence of subsets of $\mathbb{R}^n$. Suppose that*

$$E \subset \bigcup_{k=1}^{\infty} E_k.$$

Then

$$\mathcal{L}^n(E) \leq \sum_{k=1}^{\infty} \mathcal{L}^n(E_k).$$

Corollary 10.4 Let $E, E_1, E_2, E_3, \dots$ be a sequence of subsets of $\mathbb{R}^n$. Suppose that

$$E \subset \bigcup_{k=1}^{\infty} E_k$$

and that f is a nonnegative function defined at least on the set $\bigcup_{k=1}^{\infty} E_k$. Then

$$\overline{\int_E f(x) dx} \leq \sum_{k=1}^{\infty} \overline{\int_{E_k} f(x) dx}.$$

Exercise 810 Show, for all intervals I in $\mathbb{R}^n$, that $\mathcal{L}^n(I) = |I|$. □

Exercise 811 Let f and g be nonnegative functions on a set $E \subset \mathbb{R}^n$ and such that $f(x) \leq g(x)$ for all $x \in E$. Show that

$$\overline{\int_E f(x) dx} \leq \overline{\int_E g(x) dx}.$$

□

Exercise 812 Let f be a nonnegative function on a set $E \subset \mathbb{R}^n$ and such that $r \leq f(x) \leq s$ for all $x \in E$ for some real numbers r and s . Show that

$$r \mathcal{L}^n(E) \leq \overline{\int_E f(x) dx} \leq s \mathcal{L}^n(E)$$

□

Exercise 813 Suppose that $E_1, E_2 \subset \mathbb{R}^n$ are separated by open sets, i.e., there is a disjoint pair of open sets G_1 and G_2 in $\mathbb{R}^n$ so that $E_1 \subset G_1$ and $E_2 \subset G_2$. Show that

$$\overline{\int_{E_1 \cup E_2} f(x) dx} = \overline{\int_{E_1} f(x) dx} + \overline{\int_{E_2} f(x) dx}.$$

□

Exercise 814 Suppose that $E_1, E_2 \subset \mathbb{R}^n$ are separated, i.e.,

$$\inf\{\|e_1 - e_2\| : e_1 \in E_1, e_2 \in E_2\} > 0.$$

Show that

$$\overline{\int_{E_1 \cup E_2} f(x) dx} = \overline{\int_{E_1} f(x) dx} + \overline{\int_{E_2} f(x) dx}.$$

□

Exercise 815 Suppose that $E_1, E_2 \subset \mathbb{R}^n$ are separated by open sets, i.e., there is a disjoint pair of open sets G_1 and G_2 in $\mathbb{R}^n$ so that $E_1 \subset G_1$ and $E_2 \subset G_2$. Show that

$$\mathcal{L}^n(E_1 \cup E_2) = \mathcal{L}^n(E_1) + \mathcal{L}^n(E_2).$$

□

Exercise 816 Show that

$$\overline{\int_E f(x) dx} = 0$$

if and only if $f(x)$ is equal to zero for $\mathcal{L}^n$ -almost every x in E .

□

Exercise 817 Show that

$$\overline{\int_{E \cup N} f(x) dx} = 0$$

for any set N for which $\mathcal{L}^n(N) = 0$.

□

10.3 Measurable sets and measurable functions

For the definition of measurability we can repeat our theory from Chapter 7. We could choose to generalize to higher dimensions by taking any one of the characterizations of Corollary 7.24 and apply it in this setting. We choose here to take the simplest definition.

Later on in Section 10.4 we take another of the six characterizations of measurability in dimension one proved in that corollary.

Definition 10.5 A subset E of $\mathbb{R}^n$ is said to be $\mathcal{L}^n$ -measurable if for every $\varepsilon > 0$ there is an open set G with $\mathcal{L}^n(G) < \varepsilon$ and so that $E \setminus G$ is closed.

With only minor changes in wording we can prove, using the methods of Chapter 7, that the usual properties of one-dimensional Lebesgue measure are enjoyed also by $\mathcal{L}^n$. Here is a fast summary.

- Let $E_1, E_2, E_3, \dots$ be a sequence of pairwise disjoint $\mathcal{L}^n$ -measurable subsets of $\mathbb{R}^n$ and write $E = \bigcup_{i=1}^{\infty} E_i$. Then, for any set $A \subset \mathbb{R}^n$,

$$\mathcal{L}^n(A \cap E) = \sum_{i=1}^{\infty} \mathcal{L}^n(A \cap E_i).$$

- The class of all $\mathcal{L}^n$ -measurable subsets of $\mathbb{R}^n$ forms a Borel family that contains all closed sets and all $\mathcal{L}^n$ -measure zero sets.
- If $E_1 \subset E_2 \subset E_3 \subset \dots$ is an increasing sequence of subsets of $\mathbb{R}^n$ then

$$\mathcal{L}^n\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n \rightarrow \infty} \mathcal{L}^n(E_n).$$

10.3.1 Measurable functions

Definition 10.6 Let E be a $\mathcal{L}^n$ -measurable subset of $\mathbb{R}^n$ and f a real-valued function defined on E . Then f is said to be $\mathcal{L}^n$ -measurable if

$$\{x \in E : f(x) > r\}$$

is a $\mathcal{L}^n$ -measurable subset of $\mathbb{R}^n$ for every real number r .

Definition 10.7 Let E be a $\mathcal{L}^n$ -measurable subset of $\mathbb{R}^n$ and f a $\mathcal{L}^n$ -measurable function defined on E . Then the Lebesgue integral

$$\int_E f(x) dx$$

is be defined to be the value

$$\overline{\int_E [f(x)]^+ dx} - \overline{\int_E [f(x)]^- dx}$$

provided that both of these are not infinite. If both of these are finite then f is said to be Lebesgue integrable on E and the integral $\int_E f(x) dx$ has a finite value.

The key reason for this definition and for the restriction of the integration theory to measurable functions is the following fundamental additive property.

Theorem 10.8 Let E be a $\mathcal{L}^n$ -measurable subset of $\mathbb{R}^n$ and f, g be $\mathcal{L}^n$ -measurable functions defined on E . Then

$$\int_E (f(x) + g(x)) dx = \int_E f(x) dx + \int_E g(x) dx$$

provided that these are defined. (In particular this identity is valid if both f and g are Lebesgue integrable on E .)

Combining this additive theorem with the property of Lemma 10.2 we have immediately one of our most useful tools in the integration theory.

Theorem 10.9 Let E be a $\mathcal{L}^n$ -measurable subset of $\mathbb{R}^n$ and let $f_1, f_2, f_3, \dots$ be a sequence of nonnegative real-valued functions defined and Lebesgue integrable on E . Suppose that the series

$$f(x) = \sum_{k=1}^{\infty} f_k(x)$$

converges for every $x \in E$. Then

$$\int_E f(x) dx = \sum_{k=1}^{\infty} \int_E f_k(x) dx.$$

In particular, f is Lebesgue integrable on E if and only if the series of integrals converges.

Exercise 818 Show that, for any simple function

$$f(x) = \sum_{k=1}^n c_k \chi_{E_k}(x)$$

where $E_1, E_2, E_3, \dots, E_n$ are $\mathcal{L}^n$ -measurable, that

$$\int_E f(x) dx = \sum_{k=1}^n c_k \mathcal{L}^n(E \cap E_k).$$

Answer $\square$

Exercise 819 Show that any nonnegative $\mathcal{L}^n$ -measurable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ can be written in the form

$$f(x) = \sum_{k=1}^{\infty} c_k \chi_{E_k}(x)$$

for appropriate $\mathcal{L}^n$ -measurable sets $E_1, E_2, E_3, \dots$, and that

$$\int_E f(x) dx = \sum_{k=1}^{\infty} c_k \mathcal{L}^n(E \cap E_k).$$

Answer $\square$

Exercise 820 Suppose that $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a $\mathcal{L}^n$ -measurable function that is integrable on an interval I . Show that, for every $\varepsilon > 0$ there is a full interval cover β of I so that if $\pi \subset \beta$ is a packing with $J \subset I$ for each $(J, x) \in \pi$ then

$$\sum_{(J,x) \in \pi} \left| \int_J f(t) dt - f(x)|J| \right| < \varepsilon.$$

$\square$

10.3.2 Notation

We have preserved the notation from the elementary calculus in the expression

$$\int_E f(x) dx$$

interpreting now x as a dummy variable representing an arbitrary point in $\mathbb{R}^n$. There are other suggestive notations that assist in some situations. For example if $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ and E is a subset of $\mathbb{R}^2$ then the integral may appear instead as

$$\int \int_E f(x_1, x_2) dx_1 dx_2$$

or

$$\int \int_E f(x, y) dx dy.$$

The “double” integral $\int \int$ represents the fact that the dimension is two and contains a hint that an iterated integral may be useful in its computation (see Section 10.5 below).

10.4 General measure theory

The set function

$$E \rightarrow \overline{\int_E f(x) dx}$$

is defined for every subset E of $\mathbb{R}^n$. Such set functions play a role in many investigations and the students should be made acquainted with the usual general theory and its techniques.

Definition 10.10 A set function $\mathcal{M}$ defined for all subsets E of $\mathbb{R}^n$ is said to be a measure on $\mathbb{R}^n$ provided that

1. $\mathcal{M}(\emptyset) = 0$.
2. $0 \leq \mathcal{M}(E) \leq \infty$ for all subsets E of $\mathbb{R}^n$.
3. $\mathcal{M}(E_1) \leq \mathcal{M}(E_2)$ if $E_1 \subset E_2 \subset \mathbb{R}^n$.
4. $\mathcal{M}(\bigcup_{k=1}^{\infty} E_k) \leq \sum_{k=1}^{\infty} \mathcal{M}(E_k)$ for any sequence $\{E_k\}$ of subsets of $\mathbb{R}^n$.

If, moreover,

$$\mathcal{M}(E_1 \cup E_2) = \mathcal{M}(E_1) + \mathcal{M}(E_2).$$

whenever

$$\inf\{\|e_1 - e_2\| : e_1 \in E_1, e_2 \in E_2\} > 0$$

then $\mathcal{M}$ is said to be a metric measure on $\mathbb{R}^n$.

Note that Lebesgue measure $\mathcal{L}^n$ and the set function

$$\mathcal{M}(E) = \overline{\int_E f(x) dx}$$

for any nonnegative function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ are metric measures according to this definition. Many authors reserve the term “measure” for set functions defined only on special classes of sets and with stronger additive properties; they would then prefer the term “outer measure” for the concept introduced in this definition. In your readings this should not be hard to keep track of.

For the definition of measurability we take another one of the six characterizations of measurability in dimension one that we presented in Corollary 7.24.

Definition 10.11 A subset E of $\mathbb{R}^n$ is said to be $\mathcal{M}$ -measurable if for every set $A \subset \mathbb{R}^n$

$$\mathcal{M}(A) = \mathcal{M}(A \cap E) + \mathcal{M}(A \setminus E).$$

We can prove that this definition of measurability, applied to the Lebesgue measure is equivalent to that we are currently using in Definition 10.5. Using this new definition a more general theory emerges that applies to any measure on $\mathbb{R}^n$ (or indeed on any suitable space equipped with a measure). Here is a fast summary.

- Let $E_1, E_2, E_3, \dots$ be a sequence of pairwise disjoint $\mathcal{M}$ -measurable subsets of $\mathbb{R}^n$ and write $E = \bigcup_{i=1}^{\infty} E_i$. Then, for any set $A \subset \mathbb{R}^n$,

$$\mathcal{M}(A \cap E) = \sum_{i=1}^{\infty} \mathcal{M}(A \cap E_i).$$

- The class of all $\mathcal{M}$ -measurable subsets of $\mathbb{R}^n$ forms a Borel family that contains all $\mathcal{M}$ -measure zero sets.
- If $\mathcal{M}$ is a metric measure then the class of all $\mathcal{M}$ -measurable subsets contains all closed sets.

This material is standard and should be part of the background for any advanced student. Almost all texts that discuss outer measures will provide detailed proofs of these facts. You may wish to consult Chapters 2 and 3 of our text Bruckner, Bruckner, and Thomson, *Real Analysis*, 2nd Ed., ClassicalRealAnalysis.com (2008). Those chapters are available for free download.

10.5 Iterated integrals

In many cases the computation of a integral in a higher dimensional space can be accomplished only through a series of one-dimensional integrations. We do not have anything that is as convenient and useful as the calculus computation

$$\int_a^b F'(x) dx = F(b) - F(a)$$

that did most of the work in our first calculus course. But if we can reduce an integral in $\mathbb{R}^n$ to several ordinary integrals then the computations can be carried out.

The reader has likely seen in some elementary calculus classes the computation

$$\int \int_{[a,b] \times [c,d]} f(x,y) dx dy = \int_c^d \left(\int_a^b f(x,y) dx \right) dy = \int_a^b \left(\int_c^d f(x,y) dy \right) dx.$$

Another similar, and no doubt also familiar, kind of computation appears in the form

$$\int \int_E f(x,y) dx dy = \int_a^b \left(\int_{L(u)}^{U(u)} f(u,v) dv \right) du$$

when E is the set

$$E = \{(u,v) \in \mathbb{R}^2 : a \leq u \leq b, L(u) \leq v \leq U(u)\}.$$

To formulate the problem correctly we need to consider how best to state it. For example, what would we wish to state for a three dimensional Lebesgue integral

$$\int \int \int_{[a,b] \times [c,d] \times [e,f]} F(x,y,z) dx dy dz?$$

We might wish to have three iterations

$$\int_a^b \left(\int_c^d \left(\int_e^f F(x,y,z) dz \right) dy \right) dx$$

performed in the order here as 3–2–1. But there are six possible orders in which we could iterate. We also might wish to iterate this as

$$\int_e^f \left(\int \int_{[a,b] \times [c,d]} F(x,y,z) dx dy \right) dz$$

in the order (1×2) –3. There are three possible such orders in which this might be performed. To capture all of these it is best to keep to a level of abstraction. This is more convenient inside the general theory of measure and integration by using product measures. We will be a little less ambitious.

10.5.1 Formulation of the iterated integral property

Let m and n be positive integers and consider the Lebesgue integral

$$\int_I f(x) dx$$

for a function $f : I \rightarrow \mathbb{R}^{m+n} \rightarrow \mathbb{R}$ where I is an interval in $\mathbb{R}^{m+n}$. Every point x in $\mathbb{R}^{m+n}$ can be written as

$$x = (u, v) \quad (u \in \mathbb{R}^m, v \in \mathbb{R}^n)$$

and the interval $I = A(1) \times A(2)$ where $A(1)$ is an interval in $\mathbb{R}^m$ and $A(2)$ is an interval in $\mathbb{R}^n$.

We shall ask for conditions on a function $f : \mathbb{R}^{m+n} \rightarrow \mathbb{R}$ that is integrable on the interval $I = A(1) \times A(2)$ so that

- For every¹ $u \in A(1)$ the function

$$v \rightarrow f(u, v)$$

is integrable over $A(2)$ and

- the function

$$u \rightarrow \int_{A(2)} f(u, v) dv$$

is integrable on $A(1)$, and

- the identity

$$\int_{A(1) \times A(2)} f(x) dx = \int \int_{A(1) \times A(2)} f(u, v) du dv = \int_{A(1)} \left(\int_{A(2)} f(u, v) dv \right) du \quad (10.1)$$

is valid.

Exercise 821 Check that formula 10.2 holds if $f(x) = \chi_I(x)$ where $I = A(1) \times A(2)$ is an interval in $\mathbb{R}^{m+n}$. □

Exercise 822 Check that formula 10.2 holds if $f(x)$ is a step function on $I = A(1) \times A(2)$ assuming values $c_1, c_2, \dots, c_k$ on subintervals $I_1, I_2, \dots, I_k$ of I . □

¹Here we insist on every u , but as we know we could and should sometimes ignore a set of measure zero where this fails. That will be covered in Section 10.5.2.

Exercise 823 Check that formula 10.2 holds if $f(x)$ is a bounded function for which there exists a sequence of step functions $S_1, S_2, S_3, \dots$ on $I = A(1) \times A(2)$ such that $f(x) = \lim_{k \rightarrow \infty} S_k(x)$ for every $x \in I$. Answer $\square$

Exercise 824 Show that if f is a continuous function on the closed interval I then Exercise 823 can be applied to verify the formula 10.2. $\square$

Exercise 825 Let f_1, f_2 , and G be continuous functions on the closed interval I and define a function

$$f(x) = \begin{cases} f_1(x) & \text{if } x \in I \text{ and } g(x) > 0, \\ f_2(x) & \text{if } x \in I \text{ and } g(x) \leq 0. \end{cases}$$

Show that Exercise 823 can be applied to verify the formula 10.2. Answer $\square$

Exercise 826 (counterexample #1) There are a number of standard counterexamples that show some caution is needed in applying the iterated technique to multiple integrals of unbounded functions. On the interval $[-1, 1] \times [-1, 1]$ in $\mathbb{R}^2$ define the function

$$f(x, y) = xy(x^2 + y^2)^{-2} \quad f(0, 0) = 0.$$

Examine the integrals

$$\int \int_{[-1, 1] \times [-1, 1]} f(x, y) \, dx \, dy,$$

$$\int_{-1}^1 \left(\int_{-1}^1 f(x, y) \, dx \right) dy,$$

and

$$\int_{-1}^1 \left(\int_{-1}^1 f(x, y) \, dy \right) dx.$$

Answer $\square$

Exercise 827 (counterexample #2) On the interval $[0, 1] \times [-1, 1]$ in $\mathbb{R}^2$ define the function

$$f(x, y) = yx^{-3} \quad \text{if } x > 0 \text{ and } -x < y < x$$

with $f(x, y) = 0$ elsewhere. Examine the integrals

$$\int \int_{[0, 1] \times [-1, 1]} f(x, y) \, dx \, dy,$$

$$\int_{-1}^1 \left(\int_0^1 f(x,y) dx \right) dy,$$

and

$$\int_0^1 \left(\int_{-1}^1 f(x,y) dy \right) dx.$$

Answer ☐

Exercise 828 (counterexample #3) On the interval $[0, 1] \times [0, 1]$ in $\mathbb{R}^2$ define the function

$$f(x,y) = 2(x-y)(x+y)^{-3} \quad (x > 0, y > 0)$$

with $f(x,y) = 0$ elsewhere. Examine the integrals

$$\int \int_{[0,1] \times [0,1]} f(x,y) dx dy,$$

$$\int_0^1 \left(\int_0^1 f(x,y) dx \right) dy,$$

and

$$\int_0^1 \left(\int_0^1 f(x,y) dy \right) dx.$$

Answer ☐

Exercise 829 A “clever” student points out that all this trouble over integrals in $\mathbb{R}^2$ (or indeed in any dimension) can easily be avoided by simply defining double integrals as being two iterated integrals. Thus instead of proving that

$$\int \int_{[a,b] \times [c,d]} f(x,y) dx dy = \int_a^b \left(\int_c^d f(x,y) dy \right) dx$$

we just take that as a definition. Any comments?

Answer ☐

10.5.2 Fubini's theorem

In the preceding section we obtained a limited version of the iterated integral property, one that applied only to bounded functions and which required in the iteration (10.2) that the inside integral

$$\int_{A(2)} f(u, v) dv$$

exist for every value of u . The most general theorem, usually described as Fubini's theorem, asserts that this iteration is available for all integrable functions provided that we accept a set of measure zero where the inside integral might not exist.

Here are the ingredients of that theorem.

Let m and n be positive integers and we suppose that $f : \mathbb{R}^{m+n} \rightarrow \mathbb{R}$ is a function Lebesgue integrable on an interval $I = A(1) \times A(2)$ where $A(1)$ is an interval in $\mathbb{R}^m$ and $A(2)$ is an interval in $\mathbb{R}^n$. As before every point x in $\mathbb{R}^{m+n}$ is to be written as

$$x = (u, v) \quad (u \in \mathbb{R}^m, v \in \mathbb{R}^n).$$

Then

- There is a set $N(1) \subset A(1)$ with m -dimensional Lebesgue measure equal to zero.
- For every $u \in A(1) \setminus N(1)$ the function

$$v \rightarrow f(u, v)$$

is integrable over $A(2)$ and

- the function

$$u \rightarrow \int_{A(2)} f(u, v) dv$$

is integrable on $A(1)$, and

- the identity

$$\int_{A(1) \times A(2)} f(x) dx = \int \int_{A(1) \times A(2)} f(u, v) du dv = \int_{A(1) \setminus N(1)} \left(\int_{A(2)} f(u, v) dv \right) du \quad (10.2)$$

is valid.

This theorem is proved as Theorem 7–1, pp. 300–303 in E. J. McShane, *Unified Integration*, Academic Press (1983). There is a version in Chapter 6 of R. Henstock, *Lectures on the Theory of Integration*, World Scientific (1988). His version is more general (and less accessible) but uses the same defining structure essentially. The reader is, however, encouraged now to learn this theorem in the setting of general measure theory where the arguments are simpler and more straightforward. For that there are an abundance of excellent texts. We cannot resist recommending, from among them, Bruckner, Bruckner, and Thomson, *Real Analysis*, 2nd Ed., ClassicalRealAnalysis.com (2008).

10.6 Expression as a Stieltjes integral

Suppose that $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a $\mathcal{L}^n$ -measurable function that is integrable on a measurable set E . We shall show that the Lebesgue integral

$$\int_E f(x) dx$$

can be realized as a one-dimensional Stieltjes integral. Let us fix f and E for our discussion in this section and suppose that $\mathcal{L}^n(E) < \infty$. We define for each real number s the function

$$w(s) = \mathcal{L}^n(\{x \in E : f(x) > s\})$$

called the *distribution function* of the function f on the set E .

Then the following properties of the distribution function are easily established:

- The function $w : \mathbb{R} \rightarrow [0, \infty)$ is nonincreasing with

$$\lim_{s \rightarrow \infty} w(s) = 0 \text{ and } \lim_{s \rightarrow -\infty} w(s) = \mathcal{L}^n(E).$$

- $\mathcal{L}^n(\{x \in E : a < f(x) \leq b\}) = w(b) - w(a)$.
- $w(s+) = w(s)$ (i.e., w is continuous on the right at each point).
- $w(s-) = \mathcal{L}^n(\{x \in E : f(x) \geq s\})$.

The representation theorem expresses the Lebesgue integral of f in terms of the Stieltjes integral

$$\int_a^b s dw(s).$$

We know from our study of the Stieltjes integral that this must exist since w is a nonincreasing function.

Theorem 10.12 Suppose that $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a $\mathcal{L}^n$ -measurable function that is integrable on a measurable set E and that $\mathcal{L}^n(E) < \infty$. Then

$$\int_{\{x \in E : a < f(x) \leq b\}} f(x) dx = \int_a^b s dw(s) \quad (10.3)$$

and

$$\int_E f(x) dx = \int_{-\infty}^{\infty} s dw(s). \quad (10.4)$$

There are numerous textbooks where the details of this development can be found. A most readable account appears in pp. 76–79 of Wheeden and Zygmund, *Measure and Integral*, Marcel Dekker (1977).

Exercise 830 Prove the identity (10.3) in Theorem 10.12:

$$\int_{\{x \in E : a < f(x) \leq b\}} f(x) dx = \int_a^b s dw(s).$$

Answer ☐

Exercise 831 Deduce the identity (10.4) from the identity (10.3) in Theorem 10.12.

Answer ☐

Chapter 11

Appendix

11.1 Glossary of terms

In this section we present a fast account of some of the language of the course. These definitions are meant to refresh your memory or orient you in the right direction in your reading. It is still necessary to study the exact definitions and use them to prove theorems or to solve problems.

11.1.1 absolute continuity

In Chapter 4 of the text we discuss two notions of **absolute continuity**. We can consider that these are similar in some respects to the separate notions of pointwise continuity and uniform continuity.

The strongest version is due to Vitali and, accordingly, in this text we name it after him:

A function $F : [a, b] \rightarrow \mathbb{R}$ is **absolutely continuous in Vitali's sense** on $[a, b]$ provided that for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$\sum_{i=1}^n |F(x_i) - F(y_i)| < \varepsilon$$

whenever $\{[x_i, y_i]\}$ are nonoverlapping subintervals of $[a, b]$ for which

$$\sum_{i=1}^n (y_i - x_i) < \delta.$$

A weaker version of absolute continuity employs the concepts of zero variation and sets of measure zero:

A function $F : (a, b) \rightarrow \mathbb{R}$ is said to be **absolutely continuous** on the open interval (a, b) if F has **zero variation** on every subset N of the interval that has **measure zero**.

In more advanced courses the weaker version is more often used rather than an ϵ, δ version. A measure is absolutely continuous when it has zero value on sets of measure zero. The equivalence with the ϵ, δ version happens only in some cases. Our use in Chapter 4 is completely analogous to the modern use. Classical textbooks use only the Vitali definition.

11.1.2 absolute convergence

A series $\sum_{k=1}^{\infty} a_k$ is said to **converge absolutely** if both of the series

$$\sum_{k=1}^{\infty} a_k \quad \text{and} \quad \sum_{k=1}^{\infty} |a_k|$$

converge. If $\sum_{k=1}^{\infty} a_k$ converges but $\sum_{k=1}^{\infty} |a_k|$ diverges we say that the series **converges nonabsolutely** [or conditionally in some presentations].

The reason for the distinction is that when a series converges absolutely there is much more that one can do with it. Nonabsolutely convergent series are rather fragile; for example you cannot rearrange the terms of the series without possibly changing the sum.

This same language of convergent and absolutely convergent is used for infinite integrals. Thus we say that the integral $\int_a^{\infty} f(x) dx$ is **absolutely convergent** if both of the integrals

$$\int_a^{\infty} f(x) dx \quad \text{and} \quad \int_a^{\infty} |f(x)| dx$$

are convergent. **Nonabsolutely convergent** integrals are also rather fragile.

11.1.3 absolute convergence test

In order to **test** for the convergence of a series $\sum_{k=1}^{\infty} a_k$ it is often sufficient just to check that the corresponding series of absolute values

$$\sum_{k=1}^{\infty} |a_k|$$

converges. When the latter series converges the former series must converge.

Make sure you are familiar with the language of **absolute convergence** and **nonabsolute convergence**.

11.1.4 absolute integration

An integration method is an **absolute integration** method if whenever a function f is integrable on an interval $[a, b]$ then the absolute value $|f|$ is also integrable there. We know that the calculus integral is not an absolute integration method since we were able to find an integrable function f so that $|f|$ failed to be integrable.

The integrals of Riemann and Lebesgue are both absolute integration methods; the calculus integral and the Henstock-Kurweil integrals are nonabsolute integration methods.

11.1.5 almost everywhere

The phrase **almost everywhere** means except on a set of measure zero. For example, a function is continuous almost everywhere if the set of points where it is not continuous form a set of measure zero.

It is useful to extend this language to weaker situations:

mostly everywhere A statement holds *mostly everywhere* if it holds everywhere with the exception of a finite set of points $c_1, c_2, c_3, \dots, c_n$.

nearly everywhere A statement holds *nearly everywhere* if it holds everywhere with the exception of a sequence of points $c_1, c_2, c_3, \dots$

almost everywhere A statement holds *almost everywhere* if it holds everywhere with the exception of a set of measure zero.

11.1.6 Baire category theorem

See also **meager**.

Students carrying on to Chapter 9 will need to understand this theorem. Here is a full exposition suitable for most courses of instruction.

Portions If E is a closed set and (a, b) an open interval then

$$E \cap (a, b)$$

is called a *portion* of E provided only that $E \cap (a, b) \neq \emptyset$. It is possible that a portion could be trivial in that $E \cap (a, b)$ might contain only a single point of E ; such a point is said to be an *isolated point* of E and we should be alert to the possibility that a portion might merely contain such a point.

Baire-Osgood Theorem Our interest is in situations where $E, E_1, E_2, E_3, \dots$ is a sequence of closed sets and we wish to be assured that one of the sets E_n contains a portion of E . This requires a compactness argument; the nested interval property is particularly suited to this problem.

Exercise 832 Suppose that E and E_1 are nonempty closed sets and that E_1 contains no portion of E . Then there must exist a portion

$$E \cap (a, b)$$

so that $E_1 \cap (a, b) = \emptyset$.

Answer $\square$

Exercise 833 Suppose that $E, E_1, E_2, \dots, E_n$ are nonempty closed sets and that

$$E \subset \bigcup_{k=1}^n E_k.$$

Show that at least one of the sets E_k must contain a portion of E .

Answer $\square$

The Baire-Osgood theorem, one of the basic tools in our analysis later on, takes this exercise and extends the result to infinite sequences of closed sets.

Exercise 834 (Baire-Osgood Theorem) Suppose that $E, E_1, E_2, \dots, E_n, \dots$ are nonempty closed sets and that

$$E \subset \bigcup_{k=1}^{\infty} E_k.$$

Then at least one of the sets E_k must contain a portion of E .

Answer $\square$

Exercise 835 Later on we will need this theorem without having to assume that E is closed. Show that theorem remains true if $E = \bigcap_{j=1}^{\infty} G_j$ where $\{G_j\}$ is some sequence of open sets.

Answer $\square$

Exercise 836 If the closed set E is contained in a sequence of sets $\{E_n\}$ but we cannot be assured that they are closed sets then a simple device is to replace them by their closures. [The closure of a set E is the set $\bar{E}$ defined as the smallest closed set containing E .] If we do this show that the conclusion of the theorem would have to be, not that some set E_n contains a portion of E , but that some set E_n is dense in a portion of E . $\square$

Language of meager/residual subsets The exploitation of the Osgood-Baire theorem can often be clarified by using the language of meager and residual subsets. If E is a closed set¹ of real numbers then a meager subset is one that represents a “small,” insubstantial part of E ; what remains after a meager subset is removed would be called a residual subset. It would be considered a “large” subset since only an insubstantial part has been removed. Residual sets are dense, but more than dense. A countable intersection of residual sets would still be dense.

Definition 11.1 Let E be a closed set. A subset A of E is said to be a meager subset of E provided that there exists a sequence of closed sets $\{E_n\}$ none of which contains a portion of E so that

$$A \subset \bigcup_{n=1}^{\infty} E_n.$$

Definition 11.2 Let E be a closed set. A subset A of E is said to be a residual subset of E provided that the complementary subset $E \setminus A$ is a meager subset of E .

¹In this section the language is restricted to subsets of closed sets. In view of Exercise 835 all of this would apply equally well to subsets of $\mathcal{G}_\delta$ sets, that is sets that are intersections of some sequence of open sets.

11.1.7 Bolzano-Weierstrass argument

While a sequence $\{s_n\}$ may not be convergent there are many situations in which there is a convergent subsequence $\{s_{n_k}\}$. The **Bolzano-Weierstrass theorem** asserts that every bounded sequence must have at least one convergent subsequence.

This is particularly easy to prove if you first notice that all sequences have monotone subsequences.

11.1.8 bounded set

A nonempty set of real numbers is **bounded** if there is a real number M so that $|x| \leq M$ for all x in the set.

More often we would split this into **upper bounds** and **lower bounds** by finding two numbers m and M so that

$$m \leq x \leq M$$

for all x in the set. If we can find only M then we would say that the set is bounded above. If we can find only m then we would say that the set is bounded below.

If E is a bounded set then you should be able to find M and m so that

$$m \leq x \leq M \quad \text{for all } x \in E.$$

What are the best numbers for this inequality. It makes little sense, if we want to be precise, to take just any m and M that happen to work. There “must be” a maximum choice for m and a minimum choice for M . Those choices are called the **infimum** and **supremum** and abbreviated as $\inf E$ and $\sup E$. It is a deep property of the real numbers that such points do exist.

Thus we write for a bounded set E ,

$$\inf E \leq x \leq \sup E \quad \text{for all } x \in E.$$

Note 1. If E is not bounded above then we use the symbol $\sup E = \infty$ to indicate that. If E is not bounded below then we use the symbol $\inf E = -\infty$ to indicate that.

Note 2. What if $E = \emptyset$ is empty? Is it bounded? Does it have a sup and inf? We usually agree that empty sets are bounded and we commonly write the (rather mysterious) expressions $\sup \emptyset = -\infty$ and $\inf \emptyset = \infty$. Just take this as the convention and don’t fuss too much about the meaning. If precise definitions are given then one would have to conclude from those definitions that this convention is valid.

Note 3. A **sup** is also called a **least upper bound**. An **inf** is also called a **greatest lower bound**. Note that this is accurate: among all the upper bounds the minimum one is the sup. Similarly among all the lower bounds the maximum one is the inf.

11.1.9 bounded function

A function f is **bounded** if the set of values assumed by the function is bounded, i.e., if there is a number M so that $|f(x)| \leq M$ for all x in the domain of the function.

11.1.10 bounded sequence

A sequence $\{s_n\}$ is **bounded** if the set of values assumed by the sequence is bounded, i.e., if there is a number M so that $|s_n| \leq M$ for all $n = 1, 2, 3, \dots$

11.1.11 bounded variation

A function $F : [a, b] \rightarrow \mathbb{R}$ is said to be of **bounded variation** if there is a number M so that

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| \leq M$$

for all choices of points

$$a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b.$$

The least such number M is called the **total variation** of F on $[a, b]$ and is written $V(F, [a, b])$. If F is not of bounded variation then we set $V(F, [a, b]) = \infty$.

11.1.12 bounded monotone sequence argument

While a bounded sequence $\{s_n\}$ need not converge and while a monotone sequence $\{s_n\}$ need not converge, if the sequence is both **bounded and monotone** then it converges: either

$$s_1 \leq s_2 \leq s_3 \leq s_4 \leq \dots \leq s_n \nearrow L$$

or

$$s_1 \geq s_2 \geq s_3 \geq s_4 \geq \cdots \geq s_n \searrow L.$$

If the sequence is unbounded then it does not converge and, in fact, either

$$s_1 \leq s_2 \leq s_3 \leq s_4 \leq \cdots \leq s_n \nearrow \infty$$

or

$$s_1 \geq s_2 \geq s_3 \geq s_4 \geq \cdots \geq s_n \searrow -\infty.$$

11.1.13 Cantor dust

This is the **Cantor set**. See the material in Chapter 4 for a construction. It is the most important example of a closed set without isolated points, that contains no interval [i.e., it is nowhere dense], that is uncountable and that is a set of measure zero.

11.1.14 Cauchy sequences

A sequence $\{s_n\}$ converges to a number L if for $\varepsilon > 0$ there is an integer N large enough so that

$$|s_n - L| < \varepsilon$$

whenever $n \geq N$. If we are required to prove convergence of a sequence using the definition we would need to know the limit value L in advance and to have some intimate knowledge of how it relates to s_n otherwise it will be difficult to work with the definition.

The **Cauchy criterion** for convergence allows us to bypass any need for the actual limit L :

A sequence $\{s_n\}$ converges if and only if for every $\varepsilon > 0$ there is an integer N large enough so that

$$|s_n - s_m| < \varepsilon$$

whenever $n, m \geq N$.

Sequences with this property are said to be **Cauchy sequences**. The language is a bit odd since, evidently, convergent sequences are Cauchy sequences and Cauchy sequences are convergent sequences. The reason why the language survives is that in other settings than the real numbers there is an important distinction between the two ideas.

11.1.15 characteristic function of a set

Let $E \subset \mathbb{R}$. Then a convenient function for discussing properties of the set E is the function

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \text{ is in the set } E, \\ 0 & \text{if } x \text{ is not in the set } E. \end{cases}$$

Thus this function assumes only the values 0 and 1 and is defined to be 1 on E and to be 0 at every other point. This is called the **characteristic function** of E or, sometimes, the **indicator function**.

11.1.16 closed set

A set is said to be **closed** if the complement of that set is open. Thus, according to this definition, saying that a set F is closed says that all points outside of F are at some distance from F .

Specifically, see the definition of open set on page 450. According to that definition, for every point x that is not in F it is possible to find a positive number $\delta(x)$ so that the interval

$$(x - \delta(x), x + \delta(x))$$

contains no point in F .

Fix a point x that is not in F . Then, not only is there some open interval that contains x and has no points belonging to F , there is a largest such interval, say (c, d) . That interval is called a component of the open set $\mathbb{R} \setminus F$. If c [or d] happens to be finite then that point would have to belong to F (otherwise we didn't make (c, d) as large as possible).

Sometimes, when both c and d are finite, the interval $[c, d]$ is said to be *contiguous* to F . Both endpoints c and d belong to F but no point inside from (c, d) can belong to F .

11.1.17 compactness argument

By a compactness argument in the calculus is meant the invoking of one of the following arguments: the Cousin covering argument, the Bolzano-Weierstrass argument or the Heine-Borel argument.

All of these are particularly adapted to handling analysis on closed, bounded sets. Since closed, bounded sets are said to be *compact* these arguments are classified as compactness arguments. There are versions of compactness arguments in many other parts of analysis.

11.1.18 connected set

A set of real numbers E is **disconnected** if it is possible to find two disjoint open sets G_1 and G_2 so that both sets contain at least one point of E and together they include all of E . Otherwise a set is **connected**.

11.1.19 convergence of a sequence

A sequence $\{s_n\}$ **converges** to a number L if the sequence values s_n are close to L and remain close to L for large enough integers n . The formal definition must be stated in the usual ε, N form.

A sequence $\{s_n\}$ converges to a number L if for every $\varepsilon > 0$ there is an integer N large enough so that

$$|s_n - L| < \varepsilon$$

whenever $n \geq N$.

11.1.20 component of an open set

A set G is an **open set** if every point is contained in an open interval (c, d) that is itself contained in G . Thus every point is inside some open interval, where (if you like) it resides. In fact the set G can be show to be nothing but a collection of such open intervals where the points “reside.” Thus either

$$G = \emptyset \quad \text{or} \quad G = \bigcup_{k=1}^n (c_k, d_k) \quad \text{or} \quad G = \bigcup_{k=1}^{\infty} (c_k, d_k)$$

where the open intervals (c_k, d_k) are disjoint. These intervals are called the **components** of G . Thus a set can have no components [the empty set], finitely many components or a sequence of components.

The component containing a particular point x_0 in G can be accurately described as the *largest* open interval (c, d) that contains x_0 and is contained inside G .

11.1.21 composition of functions

Suppose that f and g are two functions. For some values of x it is possible that the application of the two functions one after another

$$f(g(x))$$

has a meaning. If so this new value is denoted $f \circ g(x)$ or $(f \circ g)(x)$ and the function is called the *composition* of f and g . The domain of $f \circ g$ is the set of all values of x for which $g(x)$ has a meaning and for which then also $f(g(x))$ has a meaning; that is, the domain of $f \circ g$ is

$$\{x : x \in \text{dom}(g) \text{ and } g(x) \in \text{dom}(f)\}.$$

Note that the order matters here so $f \circ g$ and $g \circ f$ have, usually, radically different meanings. This is likely one of the earliest appearances of an operation in elementary mathematics that is not commutative and that requires some care.

11.1.22 constant of integration

This is part of the theory of the indefinite integral. The symbol

$$\int f(x) dx$$

is meant to include all functions F whose derivative is $F'(x) = f(x)$ on some interval. The theory says that if you are able to find one such function $F(x)$ then every other such function is equivalent to $F(x) + C$ for some constant C . Thus we may consider that we have a *formula* for indefinite integrals:

$$\int f(x) dx = F(x) + C.$$

Here $F(x)$ is any one of the many possible functions whose derivative is $f(x)$ and C is interpreted as a completely arbitrary constant, called the **constant of integration**.

11.1.23 continuous function

For beginning calculus courses the term continuity refers simply to the property

$$\lim_{x \rightarrow x_0} f(x) = f(x_0)$$

that a function might have at a point x_0 . For the purposes of this course you should study carefully the notions of **pointwise continuity** and **uniform continuity** that appear in Chapter 1.

11.1.24 contraposition

The most common mathematical assertions that we wish to prove can be written symbolically as

$$\mathbf{P} \implies \mathbf{Q},$$

which we read aloud as “Statement **P** implies statement **Q**.” The real meaning attached to this is simply that if statement **P** is true, then statement **Q** is true.

A moment’s reflection about the meaning shows that the two versions

If **P** is true, then **Q** must be true

and

If **Q** is false, then **P** must be false

are identical in meaning. These are called **contrapositives** of each other. Any statement

$$\mathbf{P} \implies \mathbf{Q}$$

has a contrapositive

$$\text{not } \mathbf{Q} \implies \text{not } \mathbf{P}$$

that is equivalent. To prove a statement it is sometimes better not to prove it directly, but instead to prove the contrapositive.

11.1.25 converse

Suppose that we have just completed, successfully, the proof of a theorem expressed symbolically as

$$\mathbf{P} \implies \mathbf{Q}.$$

A natural question is whether the converse is also true. The **converse** is the opposite implication

$$\mathbf{Q} \implies \mathbf{P}.$$

Indeed once we have proved any theorem it is nearly routine to ask if the converse is true. Many converses are false, and a proof usually consists in looking for a counterexample.

11.1.26 countable set

That a set of real numbers is **countable** only means that it is possible to describe a sequence of real numbers

$$s_1, s_2, s_3, \dots$$

that contains every element of the set. This definition seems harmless enough but it has profound consequences and surprising conclusions.

11.1.27 Cousin's partitioning argument

Suppose that $[a, b]$ is a closed, bounded interval and at every point x of that interval we have been provided with a positive number $\delta(x)$. Then we can manufacture a partition of the interval $[a, b]$ using small intervals, intervals whose smallness is measured by $\delta(x)$.

From $[a, b]$ we are able to choose points

$$a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$$

arriving at a collection of intervals

$$\{[x_{i-1}, x_i] : i = 1, 2, 3, \dots, n\}$$

forming a finite number of nonoverlapping intervals whose union is the whole interval $[a, b]$ in such a way that we can associate some appropriate point ξ_i to each of these intervals $[x_{i-1}, x_i]$.

We can do this in such a way that the collection

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

satisfies the requirement that

$$x_i - x_{i-1} < \delta(\xi_i).$$

To use the argument, just state this:

Let $\delta(x) > 0$ for each $a \leq x \leq b$. Then there must exist a partition

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of the whole interval $[a, b]$ such that

$$\xi_i \in [x_{i-1}, x_i] \text{ and } x_i - x_{i-1} < \delta(\xi_i).$$

The argument will prove useful when you need to take a local property [here expressed by $\delta(x) > 0$] and deduce a global property [here expressed by the partition of the interval].

11.1.28 Cousin's covering argument

Cousin's partitioning argument was expressed in the language of local smallness, i.e., at each point x of an interval $[a, b]$ we were provided with a small, positive number $\delta(x)$. Using that we could construct a partition

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of the interval $[a, b]$ with each interval $[x_{i-1}, x_i]$ having length smaller than $\delta(\xi_i)$.

We can translate this into the language of **covering relations** and gain some flexibility as well as prepare us for deeper covering arguments. The collection of interval-point pairs

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

is an example of a covering relation. A *relation* because there is an association between the interval $[x_{i-1}, x_i]$ and the corresponding points ξ_i ; a *covering* because the points are invariably inside the interval. We say any collection C at all is **covering relation** if it contains only pairs $([x, y], \xi)$ where $[x, y]$ is a closed, bounded interval and ξ is a point in $[x, y]$.

The following two statements translate the Cousin lemma into a **covering argument**.

A covering relation C is said to be a full cover of a set E if for every $\xi \in E$ there is a $\delta > 0$ so that $([x, y], \xi) \in C$ for all $0 < y - x < \delta$.

*If C is a **full cover** of a closed, bounded interval $[a, b]$ then there exists a partition*

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of the whole interval $[a, b]$ that is a subset of C .

In many instances a covering argument is preferable to an argument that uses a small $\delta(x)$ at each point. Construct the cover $\mathcal{C}$. Check the full property at each point. Deduce the existence of a partition that can be extracted from the cover.

11.1.29 Darboux property

The **Darboux property** (also known as the **intermediate value property**) is the assertion that a function defined on an interval I has the property that, whenever x and y are points in I and c is any number between $f(x)$ and $f(y)$ there must be at least one point ξ between x and y where $f(\xi) = c$.

In particular note that such a function has this property: if there are points where f is positive and points where f is negative, then in between these points the function has a zero.

11.1.30 definite integral

In this text the definite integral of a function f on a closed, bounded interval is a number defined as

$$\int_a^b f(x) dx = F(b) - F(a)$$

where F is a suitably chosen antiderivative of f . There are a number of possible interpretations of this statement.

11.1.31 De Morgan's Laws

Many manipulations of sets require two or more operations to be performed together. The simplest cases that should perhaps be memorized are

$$A \setminus (B_1 \cup B_2) = (A \setminus B_1) \cap (A \setminus B_2)$$

and a symmetrical version

$$A \setminus (B_1 \cap B_2) = (A \setminus B_1) \cup (A \setminus B_2).$$

If you sketch some pictures these two rules become evident. There is nothing special that requires these “laws” to be restricted to two sets B_1 and B_2 . Indeed any family of sets $\{B_i : i \in I\}$ taken over any indexing set I must obey the same

laws:

$$A \setminus \left(\bigcup_{i \in I} B_i \right) = \bigcap_{i \in I} (A \setminus B_i)$$

and

$$A \setminus \left(\bigcap_{i \in I} B_i \right) = \bigcup_{i \in I} (A \setminus B_i).$$

Here $\bigcup_{i \in I} B_i$ is just the set formed by combining all the elements of the sets B_i into one big set (i.e., forming a large union). Similarly, $\bigcap_{i \in I} B_i$ is the set of points that are in all of the sets B_i , that is, their common intersection.

11.1.32 dense

A set of real numbers E is **dense** in an interval I if every subinterval of I contains a point of the set E . The most familiar example is the set of all rational numbers which is dense in $(-\infty, \infty)$ because every interval (c, d) contains a rational number.

11.1.33 derivative

For elementary calculus courses the term **derivative** refers simply to the computation

$$\lim_{x \rightarrow x_0} \frac{F(x) - F(x_0)}{x - x_0} = F'(x_0)$$

that is possible for many functions at a point x_0 . Those functions are said to be **differentiable**.

You will need a rather broad understanding of derivatives in order to proceed to study the calculus integral. All the necessary background material is reviewed in Chapter 1.

11.1.34 Devil's staircase

This is the name often given to the **Cantor function**. There is a full account of the Cantor function in Chapter 4.

11.1.35 domain of a function

The set of points at which a function is defined is called the **domain** of the function. It is an essential ingredient of the definition of any function. It should be considered incorrect to write

Let the function f be defined by $f(x) = \sqrt{x}$.

Instead we should say

Let the function f be defined with domain $[0, \infty)$ by $f(x) = \sqrt{x}$.

The first assertion is sloppy; it requires you to guess at the domain of the function. Calculus courses, however, often make this requirement, leaving it to you to figure out from a formula what domain should be assigned to the function. Often we, too, will require that you do this.

11.1.36 empty set

We use $\emptyset$ to represent the set that contains no elements, the empty set.

11.1.37 equivalence relation

A relation $x \sim y$ on a set S is said to be an equivalence relation if

1. $x \sim x$ for all $x \in S$.
2. $x \sim y$ implies that $y \sim x$.
3. $x \sim y$ and $y \sim z$ imply that $x \sim z$.

11.1.38 graph of a function

If f is a real function defined on an interval $[a, b]$ then the set of ordered pairs

$$\{(x, y) : a \leq x \leq b, y = f(x)\}$$

is called the **graph of the function**. In many presentations the graph is considered to be the function since there is nothing about the function that is not completely described by presenting its graph. In calculus courses one usually makes a distinction between a function and its graph, even though such a distinction is not particularly real.

11.1.39 partition

If $[a, b]$ is a closed bounded interval and we choose points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

then the collection of intervals

$$\{[x_{i-1}, x_i] : i = 1, 2, 3, \dots, n\}$$

forms a finite collection of nonoverlapping intervals whose union is the whole interval $[a, b]$. Many authors call this collection a partition.

When working with Riemann sums [as we do very frequently] one must associate some point ξ_i to each of these intervals $[x_{i-1}, x_i]$. Thus, for us a **partition** is actually the collection

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of interval-point pairs with the associated point as the second entry in the pair $([x_{i-1}, x_i], \xi_i)$.

11.1.40 Henstock-Kurzweil integral

In this course we address several theories of integration. They are

1. The Riemann integral.
2. The calculus integral [finite exceptional set case].
3. The calculus integral [countable exceptional set case].
4. The calculus integral [measure zero exceptional set case].
5. The Lebesgue integral.
6. The Henstock-Kurzweil integral.

The first three of these should be considered teaching integrals, designed to introduce beginning students to the theory of integration on the real line. Chapters 2 and 3 concern the calculus integral [finite exceptional set case]. The

fourth integral in the list here appears in Chapter 4. That integral includes all the others. In fact that integral, defined as a calculus integral, is equivalent to the Henstock-Kurzweil integral and includes Lebesgue's integral.

Later on, in Part Two, we will develop the full program of Lebesgue which addresses methods of constructing the integral based on measure theory and we will investigate the Henstock-Kurzweil integral in greater detail.

11.1.41 indefinite integral

Our indefinite integrals are all defined on an open interval and require that that open interval be specified. The most severe definition would be:

Let f be a function defined on an open interval I . Then a function $F : I \rightarrow \mathbb{R}$ is said to be an **indefinite integral** of f on I if $F'(x) = f(x)$ at every point x in I .

This is the definition taken in most calculus texts (although they are usually pretty sloppy about the interval I). For this text we take a more general definition:

Let f be a function defined on an open interval I except possibly at finitely many points. Then a *pointwise continuous* function $F : I \rightarrow \mathbb{R}$ is said to be an **indefinite integral** of f on I if $F'(x) = f(x)$ at every point x in I with possibly finitely many exceptions.

An indefinite integral, in either sense, is unique up to an additive constant. Thus by writing

$$\int f(x) dx = F(x) + C \text{ on the interval } I$$

we capture all possible indefinite integrals. The number C is intended to be completely arbitrary and is called the **constant of integration**.

11.1.42 indirect proof

Many proofs in analysis are achieved as **indirect proofs**. This refers to a specific method.

The method argues as follows. I wish to prove a statement **P** is true. Either **P** is true or else **P** is false, not both. If I suppose **P** is false perhaps I can prove that then something entirely unbelievable must be true. Since that unbelievable something is not true, it follows that it cannot be the case that **P** is false. Therefore, **P** is true.

The pattern of all indirect proofs (also known as **proofs by contradiction**) follows this structure: We wish to prove statement **P** is true. Suppose, in order to obtain a contradiction, that **P** is false. This would imply the following statements. (Statements follow.) But this is impossible. It follows that **P** is true as we were required to prove.

11.1.43 infs and sups

See bounded set on page 424.

11.1.44 integers

The integers (positive integers, negative integers, and zero). This includes the natural numbers (positive integers) 1, 2, 3, 4, ..., the negative integers $-1, -2, -3, -4, \dots$, and zero.

11.1.45 integral test for series

Consider an integer N and a nonnegative, monotone decreasing function f defined on the unbounded interval $[N, \infty)$ and that is assumed to be integrable on any closed, bounded subinterval.

Then the inequality

$$\int_N^{\infty} f(x) dx \leq \sum_{n=N}^{\infty} f(n) \leq f(N) + \int_N^{\infty} f(x) dx$$

is called the **integral test** for series, comparing an infinite integral with a series. Thus the series $\sum_{n=N}^{\infty} f(n)$ converges if and only if the integral

$$\int_N^{\infty} f(x) dx$$

is finite. In particular, if the integral diverges, then the series diverges as well. Since f is a monotone decreasing function, we know that

$$f(x) \leq f(n) \quad \text{for } x \in [n, \infty)$$

and

$$f(n) \leq f(x) \quad \text{for } x \in [N, n],$$

11.1. GLOSSARY OF TERMS

hence for every $n > N$

$$\int_n^{n+1} f(x) dx \leq \int_n^{n+1} f(n) dx = f(n) = \int_{n-1}^n f(n) dx \leq \int_{n-1}^n f(x) dx.$$

Since the lower estimate is also valid for $f(N)$, we get by summation

$$\int_N^{M+1} f(x) dx \leq \sum_{n=N}^M f(n) \leq f(N) + \int_N^M f(x) dx.$$

Let $M \rightarrow \infty$,

$$\int_N^{\infty} f(x) dx \leq \sum_{n=N}^{\infty} f(n) \leq f(N) + \int_N^{\infty} f(x) dx.$$

11.1.46 intermediate value property

The **intermediate value property** (also known as the **Darboux property**) is the assertion that a function defined on an interval I has the property that, whenever x and y are points in I and c is any number between $f(x)$ and $f(y)$ there must be at least one point ξ between x and y where $f(\xi) = c$.

In particular note that such a function has this property: if there are points where f is positive and points where f is negative, then in between these points the function has a zero.

11.1.47 interval

See Section 1.2 on page 2 for a full list.

11.1.48 least upper bound argument

If a set E of real numbers is nonempty and bounded then it is possible to find numbers M and m so that

$$m \leq x \leq M$$

for all x in the set. It is a basic principle in the study of the real numbers that there is a best choice for m and M . That is to say there is a minimum choice for M and a maximum choice for m . We call M the **least upper bound** of E and write it as $\sup E$. We call m the **greatest lower bound** of E and write it as $\inf E$.

A **least upper bound argument** is simply an argument that invokes this principle. To obtain your proof, construct a set E , show E is bounded and nonempty, and then claim that there is a number $M = \sup E$. Such a number has two key properties:

1. $x \leq M$ for all $x \in E$.
2. if $t < M$ then $x > t$ for some $x \in E$.

11.1.49 induction

This method may be used to try to prove any statement about an integer n . Here are the steps:

Step 1 Verify the statement for $n = 1$.

Step 2 (The induction step) Show that whenever the statement is true for any positive integer m it is necessarily also true for the next integer $m + 1$.

Step 3 Claim that the formula holds for all n by the **principle of induction**.

11.1.50 inverse of a function

Some functions allow an **inverse**. If $f : A \rightarrow B$ is a function, there is, sometimes, a function $f^{-1} : B \rightarrow A$ that is the reverse of f in the sense that

$$f^{-1}(f(a)) = a \text{ for every } a \in A$$

and

$$f(f^{-1}(b)) = b \text{ for every } b \in B.$$

Thus f carries a to $f(a)$ and f^{-1} carries $f(a)$ back to a while f^{-1} carries b to $f^{-1}(b)$ and f carries $f^{-1}(b)$ back to b . This can happen only if f is one-to-one and onto B .

11.1.51 isolated point

A point x_0 in a set E is **isolated** if there is an open interval (a, b) that contains the point x_0 but contains no other point of E , i.e.,

$$E \cap (a, b) = \{x_0\}.$$

All points of a finite set are isolated. A countable set may or may not have an isolated point. (Intervals, of course, have no isolated points.)

A set is **perfect** if it is nonempty, closed, and contains no isolated points.

11.1.52 Jordan decomposition

Every function F of bounded variation on an interval $[a, b]$ can be written as the difference of two monotonic, nondecreasing functions:

$$F(x) = G(x) - H(x).$$

The natural way to do this is to write

$$F(x) = \left(V(F, [a, x]) + \frac{F(x)}{2} \right) - \left(V(F, [a, x]) - \frac{F(x)}{2} \right)$$

in which case this expression is called the **Jordan decomposition**. This is part of the study of functions of bounded variation that appears in Chapter 3.

11.1.53 Lebesgue integral

In this course we address several theories of integration. They are

1. The Riemann integral.
2. The calculus integral [finite exceptional set case].
3. The calculus integral [countable exceptional set case].
4. The calculus integral [measure zero exceptional set case].

5. The Lebesgue integral.
6. The Henstock-Kurzweil integral.

The first three of these should be considered teaching integrals, designed to introduce beginning students to the theory of integration on the real line. Chapters 2 and 3 concern the calculus integral [finite exceptional set case]. The extensions to the countable exceptional set case and, finally, to the measure zero exceptional set case appear in Chapter 4. The Lebesgue integral, which is considered rather too difficult for a teaching integral, appears in Chapter 4 and is included in the theory there. In the presentation here Lebesgue's integral appears as a natural generalization of the calculus integral.

Later on, in Part Two, we will develop the full program of Lebesgue which addresses methods of constructing the integral based on measure theory. If you proceed no further than Chapter 4 you have indeed learned the Lebesgue integral and seen some of its properties, but you have missed the interesting trip through measure theory that a full account of Lebesgue's methods will explore.

11.1.54 limit of a function

The **limit of a function** is defined entirely and simply by ε s and δ s. Prior to the reworking of the foundations of the calculus by Cauchy around 1820 this idea was considered powerful but mysterious.

By the statement

$$\lim_{x \rightarrow c} f(x) = L$$

we mean that for any $\varepsilon > 0$ there is an open interval $(c - \delta, c + \delta)$ so that all values of $f(x)$ for x in that interval (except possibly for $x = c$) are between $L + \varepsilon$ and $L - \varepsilon$.

The more formal presentation is this:

Let f be defined at least on the set $(a, c) \cup (c, b)$ for some $a < c < b$. Then

$$\lim_{x \rightarrow c} f(x) = L$$

if, for any $\varepsilon > 0$, there is an $\delta > 0$ so that

$$L - \varepsilon < f(x) < L + \varepsilon$$

if

$$c < x < c + \delta \text{ or } c - \delta < x < c.$$

Notice that the definition carefully excludes the value of $f(c)$ from influencing the statement. Thus $f(c)$ could be any value at all or may not even be defined. Also notice that a feature of the limit definition is that it is two-sided. The function must be defined on open intervals on either side of the point c , and those values will influence the existence of the limit.

One-sided limits just copy this definition with the obvious changes:

Let f be defined at least on the set (c, b) for some $c < b$. Then

$$f(c+) = \lim_{x \rightarrow c+} f(x) = L$$

if, for any $\varepsilon > 0$, there is an $\delta > 0$ so that

$$L - \varepsilon < f(x) < L + \varepsilon$$

if $c < x < c + \delta$.

and

Let f be defined at least on the set (a, c) for some $a < c$. Then

$$f(c-) = \lim_{x \rightarrow c-} f(x) = L$$

if, for any $\varepsilon > 0$, there is an $\delta > 0$ so that

$$L - \varepsilon < f(x) < L + \varepsilon$$

if $c - \delta < x < c$.

In working with limits it will be very convenient to have access to **sequence methods**. Here is the two-sided version; one-sided versions are similar.

Let f be defined at least on the set $(a, c) \cup (c, b)$ for some $a < c < b$. Then

$$\lim_{x \rightarrow c} f(x) = L$$

if and only if

$$\lim_{n \rightarrow \infty} f(x_n) = L$$

for every sequence $\{x_n\}$ of points from $(a, c) \cup (c, b)$ that converges to c .

These methods allow us to apply the powerful methods of Cauchy sequences, Bolzano-Weierstrass property or monotone sequence arguments in discussions of limits.

11.1.55 linear combination

If $f(x)$ and $g(x)$ are functions and r and s are real numbers then the function

$$h(x) = rf(x) + s(g(x))$$

is said to be a **linear combination** of the two functions f and g . The same phrase would be used for any finite number of functions, not just two.

11.1.56 Lipschitz function

If a function F defined on an interval I has this property for some number C ,

$$|F(x) - F(y)| \leq C|x - y| \quad \text{for all } x, y \in I,$$

then F is said to be a **Lipschitz function**. Such functions arise all the time in the study of derivatives and integrals.

If you are able to determine the smallest number C with this property then that number is called the **Lipschitz constant** for the function F on the interval I .

11.1.57 locally bounded function

A function f defined on an open interval is said to be **locally bounded** at a point x_0 in that interval if there is at least one interval (c, d) contained in I and containing x_0 such that f is bounded on (c, d) . This is the same as saying that, for some positive number δ , the function is bounded on $(x_0 - \delta, x_0 + \delta)$.

11.1.58 lower bound of a set

See bounded set on page 424.

11.1.59 managing epsilons

The student should have had some familiarity with ϵ , δ proofs prior to taking on this course. Certainly the notions of continuity, limits, and derivatives in an earlier course will have prepared the rudiments.

In such arguments here we frequently have several steps or many steps. With two steps the student is by now accustomed to splitting ϵ into two pieces

$$\epsilon = \frac{\epsilon}{2} + \frac{\epsilon}{2}.$$

Then the argument needed to show something is smaller than ϵ breaks into showing the two separate pieces are smaller than $\frac{\epsilon}{2}$.

It is only moderately more difficulty to handle infinitely many pieces in a proof. In the calculus as presented in these notes we frequently have to handle some condition described by an infinite sequence of steps. For that a very simple device is available, similar to splitting the ϵ into two or three or more pieces:

$$\epsilon = \frac{\epsilon}{2} + \frac{\epsilon}{4} + \frac{\epsilon}{8} + \frac{\epsilon}{16} + \cdots + \frac{\epsilon}{2^n} + \cdots$$

It would be reasonable to use this computation even prior to any study of sequences and series. In fact, however, an infinite series can be avoided in all the proofs anyway since all sums are finite. Thus the student never needs anything beyond the inequality

$$\frac{\epsilon}{2} + \frac{\epsilon}{4} + \frac{\epsilon}{8} + \frac{\epsilon}{16} + \cdots + \frac{\epsilon}{2^n} < \epsilon.$$

This can be proved with elementary algebra.

11.1.60 meager

A set of real numbers is countable if it can be expressed as a countable union of a sequence of finite sets. If I is an interval and E is a countable set then $I \setminus E$ is dense in I .

This generalizes to meager sets. A set of real numbers is *meager* if it can be expressed as a countable union of a sequence of nowhere dense sets. If I is an interval and E is a meager set then $I \setminus E$ is dense in I . The proof for meager sets and for countable sets is exactly the same, using the nested interval argument. For example: if E_n is a sequence of nowhere dense sets [finite sets] inside an interval I , then take any subinterval $(c, d) \subset I$. There must be a nested sequence of intervals with $[c_n, d_n] \subset I$ and $[c_n, d_n] \cap E_n = \emptyset$. There is a point that belongs to all of the intervals and that point fails to belong to

$E = \bigcup_{n=1}^{\infty} E_n$. This shows that $I \setminus E$ is dense in I .

The complement of the meager set E is said to be *residual* in I . Residual sets are dense as we have just seen. This is usually described as the Baire category theorem.

See also **Baire category theorem**.

11.1.61 mean-value theorem

In the differential calculus the **mean-value theorem** is the assertion that

$$\frac{F(b) - F(a)}{b - a} = F'(\xi)$$

at some point ξ between a and b . The most economical assumptions are that F is uniformly continuous on $[a, b]$ and differentiable on (a, b) .

The corresponding assertion for the integral is also called the mean-value theorem,

$$\frac{1}{b - a} \int_a^b F'(x) dx = F'(\xi),$$

and has the same assumptions and is the same statement expressed in different language.

11.1.62 measure zero

There are a number of definitions possible and, ultimately, we would need to check that they are equivalent. The definition that we use in Section 4.3 is stated in terms of Riemann sums over subpartitions and a small measurement $\delta(x)$ at each point of the measure zero set.

A set of real numbers N is said to be a set of **measure zero** if for every $\varepsilon > 0$ and every point $\xi \in N$ there is a $\delta(\xi) > 0$ with the following property: whenever a subpartition

$$\{[c_i, d_i], \xi_i) : i = 1, 2, \dots, n\}$$

is given with each $\xi_i \in N$ and so that

$$0 < d_i - c_i < \delta(\xi_i) \quad (i = 1, 2, \dots, n)$$

then

$$\sum_{i=1}^n (d_i - c_i) < \varepsilon.$$

An equivalent definition might be more familiar to some readers since other textbooks are likely to start here:

A set of real numbers N is said to be a set of **measure zero** if for every $\varepsilon > 0$ it is possible to find a sequence of open intervals $\{(c_k, d_k)\}$ so that every point in N appears in at least one of the intervals and $\sum_{k=1}^{\infty} (d_k - c_k) < \varepsilon$.

11.1.63 monotone subsequence argument

While a sequence $\{s_n\}$ need not be monotone it always possesses at least one monotonic subsequence: either

$$s_{n_1} \leq s_{n_2} \leq s_{n_3} \leq s_{n_4} \leq \dots$$

or

$$s_{n_1} \geq s_{n_2} \geq s_{n_3} \geq s_{n_4} \geq \dots$$

In particular if the original sequence is bounded then there is a monotonic subsequence that converges.

11.1.64 mostly everywhere

The phrase almost everywhere is in common use in all advanced analysis; it means except on a set of measure zero.

It is useful to extend this language to weaker situations. For this course especially, we have made frequent use of finite exceptional sets or countable exceptional sets. Nearly everywhere is common usage, but it is not universal. **Mostly everywhere** is peculiar to us and we don't use it, but we do recommend it would be useful in classroom discussions.

mostly everywhere A statement holds *mostly everywhere* if it holds everywhere with the possible exception of a finite set of points $c_1, c_2, c_3, \dots, c_n$.

nearly everywhere A statement holds *nearly everywhere* if it holds everywhere with the possible exception of a sequence of points $c_1, c_2, c_3, \dots$.

almost everywhere A statement holds *almost everywhere* if it holds everywhere with the possible exception of a set of measure zero.

11.1.65 natural numbers

The natural numbers (positive integers): 1, 2, 3, 4, ...

11.1.66 nearly everywhere

The phrase **almost everywhere** is in common use in all advanced analysis; it means except on a set of measure zero. Nearly everywhere is often used to indicate except on a countable set. Mostly everywhere is just a suggestion. We don't use it in the text, but we do recommend it would be useful in classroom discussions.

mostly everywhere A statement holds *mostly everywhere* if it holds everywhere with the possible exception of a finite set of points $c_1, c_2, c_3, \dots, c_n$.

nearly everywhere A statement holds *nearly everywhere* if it holds everywhere with the possible exception of a sequence of points $c_1, c_2, c_3, \dots$.

almost everywhere A statement holds *almost everywhere* if it holds everywhere with the possible exception of a set of measure zero.

11.1.67 negations of quantified statements

Here is a tip that helps in forming negatives of assertions involving quantifiers. The two quantifiers $\forall$ and $\exists$ are complementary in a certain sense. The negation of the statement "All birds fly" would be (in conventional language) "Some bird does not fly." More formally, the negation of

For all birds b , b flies

would be

There exists a bird b , b does not fly.

11.1. GLOSSARY OF TERMS

In symbols let B be the set of all birds. Then the form here is

$\forall b \in B$ “statement about b ” is true

and the negation of this is

$\exists b \in B$ “statement about b ” is *not* true.

This allows a simple device for forming negatives. The negation of a statement with $\forall$ is a statement with $\exists$ replacing it, and the negation of a statement with $\exists$ is a statement with $\forall$ replacing it.

11.1.68 nested interval argument

The nested interval argument is a way of finding a point with a particular desirable property. We first construct intervals that we are sure will contain the point we want and then shrink down to the point.

Nested interval argument: If $[a_1, b_1] \supset [a_2, b_2] \supset [a_3, b_3] \dots$ is a shrinking sequence of closed, bounded intervals with lengths decreasing to zero,

$$\lim_{n \rightarrow \infty} (b_n - a_n) = 0$$

then there is a unique point z that belongs to each of the intervals.

Note that the argument requires the intervals are both closed and bounded. The examples

$$(0, 1), (1, \frac{1}{2}), (0, \frac{1}{3}), (0, \frac{1}{4}), \dots$$

and

$$[1, \infty), [2, \infty), [3, \infty), [4, \infty), \dots$$

do not have any point in the intersection.

11.1.69 nowhere dense

A set of real numbers E is **nowhere dense** in an interval $[a, b]$ if for every subinterval $[c, d] \subset [a, b]$ it is possible to find a further subinterval $[c_1, d_1] \subset [c, d]$ that contains no points of the set E . It is easy to check that a closed set of real numbers E is nowhere dense in an interval $[a, b]$ if it contains no subinterval of $[a, b]$.

Apart from finite sets, some countable sets, and the Cantor set we have not seen many nowhere dense sets in the text.

Dense and **nowhere dense** are not exact opposites, but instead extreme opposites: a dense set appears in every subinterval and a nowhere dense set makes no appearance in most subintervals.

11.1.70 open set

The notion of an open set is built up by using the idea of an open interval. A set G is said to be **open** if for every point $x \in G$ it is possible to find an open interval (c, d) that contains the point x and is itself contained entirely inside the set G .

It is possible to give an ϵ, δ type of definition for open set. (In this case just the δ is required.) A set G is open if for each point $x \in G$ it is possible to find a positive number $\delta(x)$ so that

$$(x - \delta(x), x + \delta(x)) \subset G.$$

Fix a point $x \in G$. Then, not only is there some open interval that contains x and is contained in G , there is a largest such interval. That interval is called a component of G .

See also **component of an open set** on page 428.

11.1.71 one-to-one and onto function

If to each element b in the range of f there is precisely one element a in the domain so that $f(a) = b$, then f is said to be **one-to-one** or **injective**. We sometimes say, about the range $f(A)$ of a function, that f maps A *onto* $f(A)$. If $f : A \rightarrow B$, then f would be said to be *onto* B if B is the range of f , that is, if for every $b \in B$ there is some $a \in A$ so that $f(a) = b$. A function that is onto is sometimes said to be **surjective**. A function that is both one-to-one and onto is sometimes said to be **bijective**.

11.1.72 ordered pairs

Given two sets A and B , we often need to discuss pairs of objects (a, b) with $a \in A$ and $b \in B$. The first item of the pair is from the first set and the second item from the second. Since order matters here these are called **ordered pairs**. The set of all ordered pairs (a, b) with $a \in A$ and $b \in B$ is denoted

$$A \times B$$

and this set is called the **Cartesian product** of A and B .

11.1. GLOSSARY OF TERMS

11.1.73 oscillation of a function

Let f be a function defined at least on an interval $[c, d]$. We write

$$\omega f([c, d]) = \sup\{|f(x) - f(y)| : c \leq y < x \leq d.\}$$

The number $\omega f([c, d])$ is called the **oscillation** of the function f on the interval $[c, d]$.

11.1.74 partition

If $[a, b]$ is a closed bounded interval and we choose points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

then the collection of intervals

$$\{[x_{i-1}, x_i] : i = 1, 2, 3, \dots, n\}$$

forms a finite collection of nonoverlapping intervals whose union is the whole interval $[a, b]$.

When working with Riemann sums [as we do very frequently] one must associate some point ξ_i to each of these intervals $[x_{i-1}, x_i]$. Thus, for us a **partition** is actually the collection

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of interval-point pairs with the associated point as the second entry in the pair $([x_{i-1}, x_i], \xi_i)$.

If f is a function defined at points of $[a, b]$ the sum

$$\sum_{i=1}^n f(\xi_i)(x_i - x_{i-1})$$

is called a **Riemann sum** for the function and the partition.

But we also consider sums of the following form, loosely, as Riemann sums:

$$\sum_{i=1}^n |f(\xi_i)|(x_i - x_{i-1})$$

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})|$$

$$\sum_{i=1}^n |[F(x_i - F(x_{i-1}))] - f(\xi_i)(x_i - x_{i-1})|$$

and we continue to call these Riemann sums even if we use only part of the partition in the sum (i.e., a **subpartition**).

11.1.75 perfect set

A set is **perfect** if it is nonempty, closed, and contains no isolated points.

A point x_0 in a set E is **isolated** if there is an open interval (a, b) that contains the point x_0 but contains no other point of E , i.e.,

$$E \cap (a, b) = \{x_0\}.$$

All points of a finite set are isolated. A countable set may or may not have an isolated point. (Intervals, of course, have no isolated points.)

11.1.76 pointwise continuous function

Make sure to distinguish between **pointwise continuous** and **uniform continuous**.

What in this text is called “pointwise” continuous is usually simply called continuous. We are using this as a teaching device to emphasize as much as possible the subtle and important distinction between uniform continuity and the local version that is defined in a pointwise manner. The instructor should, no doubt, apologize to the students and ask them to strike the nonstandard word “pointwise” from all such occurrences. But this pointwise/uniform distinction will occur again in the text anyway in situations where the language is standard.

Let $f : I \rightarrow \mathbb{R}$ be a function defined on an interval I . We say that f is **uniformly continuous** if for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$|f(d) - f(c)| < \varepsilon$$

whenever c, d are points in I for which $|d - c| < \delta$.

and

Let $f : I \rightarrow \mathbb{R}$ be a function defined on an open interval I and let x_0 be a point in that interval. We say that f

is **pointwise continuous** at x_0 if for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$|f(x) - f(x_0)| < \varepsilon$$

whenever x is a point in I for which $|x - x_0| < \delta$.

11.1.77 preimage of a function

If $f : X \rightarrow Y$ and $E \subset Y$, then

$$f^{-1}(E) = \{x : f(x) = y \text{ for some } y \in E\} \subset X$$

is called the **preimage** of E under f . [There may or may not be an inverse function here; $f^{-1}(E)$ has a meaning even if there is no inverse function.]

11.1.78 quantifiers

We often encounter two phrases used repeatedly:

For all ... it is true that ...

and

There exists a ... so that it is true that ...

For example, the formula

$$(x + 1)^2 = x^2 + 2x + 1$$

is true *for all* real numbers x . *There is a* real number x such that

$$x^2 + 2x + 1 = 0$$

(indeed $x = -1$).

It is extremely useful to have a symbolic way of writing this. It is universal for mathematicians of all languages to use the symbol $\forall$ to indicate “for all” or “for every” and to use $\exists$ to indicate “there exists.” Originally these were chosen since it was easy enough for typesetters to turn the characters “A” and “E” around or upside down. These are called **quantifiers**.

11.1.79 range of a function

The set of points B in the definition of a function

$$f : D \rightarrow B$$

is sometimes called the range or co-domain of the function. Most writers do not like the term “range” for this and prefer to use the term **range** for the set

$$f(A) = \{f(x) : x \in A\} \subset B$$

that consists of the actual output values of the function f , not some larger set that merely contains all these values.

11.1.80 rational numbers

The **rational numbers**, are all the fractions m/n where m and n are integers (and $n \neq 0$). The rational numbers can also be described by infinite decimal expansions: they are precisely those expansions which are either repeating or terminating.

In a proper course of instruction on the real numbers it would be verified that all real numbers have a representation as an infinite decimal expansion and, conversely, that an infinite decimal expansion does represent some real number.

In our course we are taking it for granted that the student has seen an adequate presentation of the real numbers and can sort out, from the minimal description given here, which numbers are rational and which are not (**irrational**) in simple situations.

11.1.81 real numbers

We use the symbol $\mathbb{R}$ to denote the set of all the **real numbers**. We have also agreed to interpret the set of real numbers in the language of intervals as the set $(-\infty, \infty)$.

There is insufficient space in this short glossary to define exactly what the real numbers are. From a practical point of view the student needs mainly to understand how to do simple algebra, work with inequalities and manage the use of the supremum and infimum (sup and inf).

Our textbook

[TBB] Elementary Real Analysis, 2nd Edition, B. S. Thomson, J. B. Bruckner, A. M. Bruckner, Classical-RealAnalysis.com (2008).

starts with a chapter on this (which you can download for free). At some point in your career it would be of great benefit to see a constructive presentation of the real numbers.

11.1.82 relations

Often in mathematics we need to define a relation on a set S . Elements of S could be related by sharing some common feature or could be related by a fact of one being “larger” than another. For example, the statement $A \subset B$ is a relation on families of sets and $a < b$ a relation on a set of numbers. Fractions of integers p/q and a/b are related if they define the same number; thus we could define a relation on the collection of all fractions by $p/q \sim a/b$ if $pb = qa$.

A **relation** R on a set S then would be some way of deciding whether the statement xRy (read as x is related to y) is true. If we look closely at the form of this we see it is completely described by constructing the set

$$R = \{(x, y) : x \text{ is related to } y\}$$

of ordered pairs. Thus a relation on a set is not a new concept: It is merely a collection of ordered pairs. Let R be any set of ordered pairs of elements of S . Then $(x, y) \in R$ and xRy and “ x is related to y ” can be given the same meaning. This reduces relations to ordered pairs. In practice we usually view the relation from whatever perspective is most intuitive. [For example, the order relation on the real line $x < y$ is technically the same as the set of ordered pairs $\{(x, y) : x < y\}$ but hardly anyone thinks about the relation this way.]

11.1.83 residual

See **meager**.

11.1.84 Riemann sum

For us a **partition** of an interval $[a, b]$ is a collection

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of interval-point pairs with the intervals nonoverlapping and having union all of $[a, b]$. A subpartition is just a subset of a partition. We refer to all of the following sums as Riemann sums over either a partition or subpartition.

$$\begin{aligned} & \sum_{i=1}^n (x_i - x_{i-1}), \\ & \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}), \\ & \sum_{i=1}^n |f(\xi_i)|(x_i - x_{i-1}), \\ & \sum_{i=1}^n |F(x_i) - F(x_{i-1})|, \\ & \sum_{i=1}^n |[F(x_i) - F(x_{i-1})] - f(\xi_i)(x_i - x_{i-1})|, \\ & \sum_{i=1}^n f(\xi_i)[G(x_i) - G(x_{i-1})]. \end{aligned}$$

11.1.85 Riemann integral

In this course we address several theories of integration. They are

1. The Riemann integral.
2. The calculus integral [finite exceptional set case].
3. The calculus integral [countable exceptional set case].
4. The calculus integral [measure zero exceptional set case].
5. The Lebesgue integral.
6. The Henstock-Kurzweil integral.

The first three of these should be considered teaching integrals, designed to introduce beginning students to the theory of integration on the real line. Chapters 2 and 3 concern the calculus integral [finite exceptional set case]. The Riemann integral, which is another candidate for a teaching integral, is merely mentioned in Chapter 4 although various themes associated with that integral (notably uniform approximation by Riemann sums) appear frequently in the text.

The Riemann integral is of limited use as a teaching integral. If one spends very little time with it and uses it as a platform for a fast launch of integration theories based on pointwise approximation by Riemann sums, then (possibly) it could be justified as a teaching tool.

11.1.86 series

An infinite sum

$$a_1 + a_2 + a_3 + \cdots + a_n + \cdots$$

is called a **series**. The usual notation is

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + \cdots$$

The sum of the series is the limit, if it exists, of the **sequence of partial sums**,

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n a_k = \lim_{n \rightarrow \infty} (a_1 + a_2 + a_3 + \cdots + a_n).$$

Series are sometimes called **infinite series**, but for mathematicians the term series always means an infinite sum and the word series is not used for finite sequences or finite lists as the common dictionary definition would suggest.

A series is **convergent** if the sum exists, i.e., the limit

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n a_k$$

exists (and has a finite value). It is said to be **divergent** otherwise.

If both series

$$\sum_{k=1}^{\infty} a_k \quad \text{and} \quad \sum_{k=1}^{\infty} |a_k|$$

converge then the series is said to be **absolutely convergent**. It is possible for a series to be convergent but not absolutely

convergent; such series are said to be **nonabsolutely convergent**.

If the terms of a series are summed in a different order, e.g.,

$$a_2 + a_1 + a_5 + \cdots + a_4 + a_{15} + \cdots +$$

the new series is called a **rearrangement**. If all rearrangements of a series converge then the series is said to be **unconditionally convergent**. If the series converges but has a divergent rearrangement then the series is said to be **conditionally convergent**. It is a theorem in the study of series that absolute convergence and unconditional convergence are equivalent.

11.1.87 set-builder notation

We use the notation $\{x : x^2 + x < 0\}$ to represent the set of all real numbers x satisfying the inequality $x^2 + x < 0$. It may take some time but if you are adept at inequalities and quadratic equations you can recognize that this set is exactly the open interval $(-1, 0)$. This is a useful way of describing a set (when possible): Just describe, by an equation or an inequality, the elements that belong. In general, if $C(x)$ is some kind of assertion about an object x , then $\{x : C(x)\}$ is the set of all objects x for which $C(x)$ happens to be true. Other formulations can be used. For example,

$$\{x \in A : C(x)\}$$

describes the set of elements x that belong to the set A and for which $C(x)$ is true. The example $\{1/n : n = 1, 2, 3, \dots\}$ illustrates that a set can be obtained by performing computations on the members of another set.

11.1.88 set notation

Sets are just collections of objects. If the word “set” becomes too often repeated, you might find that words such as *collection*, *family*, or *class* are used. Thus a set of sets might become a family of sets. The statement $x \in A$ means that x is one of those numbers belonging to A . The statement $x \notin A$ means that x is *not* one of those numbers belonging to A . (The stroke through the symbol $\in$ here is a familiar device, even on road signs or no smoking signs.)

11.1. GLOSSARY OF TERMS

11.1.89 subpartition

If $[a, b]$ is a closed bounded interval and we choose points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

then the collection of intervals

$$\{[x_{i-1}, x_i] : i = 1, 2, 3, \dots, n\}$$

forms a finite collection of nonoverlapping intervals whose union is the whole interval $[a, b]$.

When working with Riemann sums [as we do very frequently] one must associate some point ξ_i to each of these intervals $[x_{i-1}, x_i]$. A **partition** is the collection

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of interval-point pairs with the associated point as the second entry in the pair $([x_{i-1}, x_i], \xi_i)$.

Any subset of a partition is a **subpartition**. Thus, in order for a collection

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

to be a subpartition the intervals must not overlap, the points ξ_i must be chosen from the corresponding interval $[x_{i-1}, x_i]$, but the union of all the intervals need not fill out an entire interval, just part of it.

11.1.90 summation by parts

The **summation by parts** formula is just the elementary identity

$$\begin{aligned} \sum_{k=m}^n a_k b_k &= \sum_{k=m}^n (s_k - s_{k-1}) b_k \\ &= s_m(b_m - b_{m+1}) + s_{m+1}(b_{m+1} - b_{m+2}) \cdots + s_{n-1}(b_{n-1} - b_n) + s_n b_n. \end{aligned}$$

It is a discrete analogue of integration by parts, and is occasionally a useful identity to have in the study of series.

11.1.91 sups and infs

See bounded set on page 424.

A **sup** is also called a **least upper bound**. An **inf** is also called a **greatest lower bound**.

11.1.92 subsets, unions, intersection, and differences

The language of sets requires some special notation that is, doubtless, familiar. If you find you need some review, take the time to learn this notation well as it will be used in all of your subsequent mathematics courses.

1. $A \subset B$ (A is a **subset** of B) if every element of A is also an element of B .
2. $A \cap B$ (the **intersection** of A and B) is the set consisting of elements of both sets.
3. $A \cup B$ (the **union** of the sets A and B) is the set consisting of elements of either set.
4. $A \setminus B$ (the difference² of the sets A and B) is the set consisting of elements belonging to A but not to B .

In the text we will need also to form unions and intersections of large families of sets, not just of two sets.

11.1.93 total variation function

Let $F : [a, b] \rightarrow \mathbb{R}$ be a function of bounded variation. Then the function

$$T(x) = V(F, [a, x]) \quad (a < x \leq b), \quad T(a) = 0$$

is called the **total variation function** for F on $[a, b]$.

See also function of **bounded variation** on page 425.

11.1.94 uniformly continuous function

Make sure to distinguish between **[pointwise] continuous** and **uniformly continuous**.

Let $f : I \rightarrow \mathbb{R}$ be a function defined on an interval I . We say that f is **uniformly continuous** if for every $\epsilon > 0$ there is a $\delta > 0$ so that

$$|f(d) - f(c)| < \epsilon$$

whenever c, d are points in I for which $|d - c| < \delta$.

²Don't use $A - B$ for set **difference** since it suggests subtraction, which is something else.

and

Let $f : I \rightarrow \mathbb{R}$ be a function defined on an open interval I and let x_0 be a point in that interval. We say that f is **pointwise continuous** at x_0 if for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$|f(x) - f(x_0)| < \varepsilon$$

whenever x is a point in I for which $|x - x_0| < \delta$.

When you are finished with the course you may drop the word “pointwise” since it is not standard. We are using this to emphasize the distinction between the two levels of continuity assumptions.

11.1.95 upper bound of a set

See bounded set on page 424.

11.1.96 variation of a function

The variation of a function concerns estimates on the size of the sums

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})|$$

for choices of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

There are two (related) definitions that we use:

A function $F : [a, b] \rightarrow \mathbb{R}$ is said to be of **bounded variation** if there is a number M so that

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| \leq M$$

for all choices of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

The least such number M is called the **total variation** of F on $[a, b]$ and is written $V(F, [a, b])$. If F is not of bounded variation then we set $V(F, [a, b]) = \infty$.

and

A function $F : (a, b) \rightarrow \mathbb{R}$ is said to have **zero variation** on a set $E \subset (a, b)$ if for every $\varepsilon > 0$ and every $x \in E$ there is a $\delta(x) > 0$

$$\sum_{i=1}^n |F(b_i) - F(a_i)| < \varepsilon$$

whenever a subpartition

$$\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$$

is chosen for which

$$\xi_i \in E \cap [a_i, b_i] \text{ and } b_i - a_i < \delta(\xi_i).$$

11.2 Answers to exercises

Exercise 1, page 4

The symbols $-\infty$ and ∞ do not stand for real numbers; they are used in various contexts to describe a *situation*. For example $\lim_{n \rightarrow \infty} \frac{n}{n+1} = 1$ and $\lim_{n \rightarrow \infty} \frac{n^2}{n+1} = \infty$ have meanings that do not depend on there being a real number called ∞ . Thus stating $a < x < \infty$ simply means that x is a real number larger than a . [It does not mean that x is a real number smaller than ∞ , because there is no such thing.]

Exercise 2, page 4

Well, we have labeled some intervals as “bounded” and some as unbounded. But the definition of a bounded set E requires that we produce a real number M so that $|x| \leq M$ for all $x \in E$ or, equivalently that $-M \leq x \leq M$ for all $x \in E$. Show that the labels are correct in terms of this definition of what bounded means.

Exercise 3, page 4

Well, we have labeled some intervals as “open” and some as not. But the definition of an open set G requires that we produce, for each $x \in G$ at least one interval (c, d) that contains x and is contained inside the set G . Show that the labels are correct in terms of the definition of what open means here. (It is almost immediate from the definition but make sure that you understand the logic and can write it down.)

Exercise 4, page 4

Again, we have labeled some intervals as “closed” and some as not. Show that the labels are correct in terms of the definition of what closed means here. Remember that the definition of closed is given in terms of the complementary set. A set E is *closed* if the set $\mathbb{R} \setminus E$ is an open set. So, for these intervals, write down explicitly what that complementary set is.

Exercise 5, page 4

For example $[a, b)$ is not open because the point a is in the set but we cannot find an open interval that contains a and is also a subset of $[a, b)$. Thus the definition fails at one point of the set. For not closed, work with the complement of $[a, b)$, i.e., the set $(-\infty, a) \cup [b, \infty)$ and find a point that illustrates that this set cannot be open.

Exercise 6, page 4

Yes, if the two open intervals have a point in common [i.e., are not disjoint]. Otherwise the intersection would be the empty set $\emptyset$; for this reason some authors [not us] call the empty set a degenerate open interval.

Exercise 7, page 4

Not in general. If the two intervals have only one point in common or no points in common the intersection is not an interval. Yes, if the two closed intervals have at least two point in common.

If we have agreed (as in the discussion to the preceding exercise) to call the empty set a degenerate open interval we would be obliged also to call it a degenerate closed interval.

Exercise 8, page 4

Not necessarily. The intersection could, of course, be the empty set which we do not interpret as an interval, and we must consider the empty set as bounded. Even if it is not empty it need not be unbounded. Consider $(\infty, 1) \cap (0, \infty) = (0, 1)$.

Exercise 9, page 4

The only possibility would be

$$(a, b) \cup (c, d) = (s, t)$$

where the two intervals (a, b) and (c, d) have a point in common. In that case $s = \min\{a, c\}$ and $t = \max\{b, d\}$. If (a, b) and (c, d) are disjoint then $(a, b) \cup (c, d)$ is not an interval, but a disjoint union of two open intervals.

Exercise 10, page 4

The only possibility would be $(-\infty, c] \cup [c, \infty) = (-\infty, \infty)$.

Exercise 11, page 4

Yes. Prove, in fact, that the union of a finite number of bounded sets is a bounded set.

Exercise 12, page 4

Remember that $A \setminus B$ is the set of all points that are in the set A but are not in the set B . If I is open you should discover that $I \setminus C$ is a union of a finite number of disjoint open intervals, and is an open set itself. For example if $I = (a, b)$ and $C = \{c_1, c_2, \dots, c_m\}$ where these are points inside (a, b) then

$$(a, b) \setminus C = (a, c_1) \cup (c_1, c_2) \cup (c_2, c_3) \cup \dots \cup (c_m, b).$$

Exercise 13, page 4

Remember that $A \setminus B$ is the set of all points that are in the set A but are not in the set B .

The set $I \setminus C$ must be a union of intervals. There are a number of possibilities and so, to answer the exercise, it is best to just catalogue them. For example if $I = [a, b]$ and $C = \{a, b\}$ then $[a, b] \setminus C$ is the open interval (a, b) . If $C = \{c_1, c_2, \dots, c_m\}$ where these are points inside (a, b) then

$$[a, b] \setminus C = [a, c_1) \cup (c_1, c_2) \cup (c_2, c_3) \cup \dots \cup (c_m, b].$$

After handling all the possibilities it should be clear that $I \setminus C$ is a union of a finite number of disjoint intervals. The intervals need not all be open or closed.

Exercise 14, page 6

If a sequence of real numbers $\{s_n\}$ converges to a real number L then, for any choice of $\epsilon_0 > 0$ there is an integer N so that

$$L - \epsilon_0 < s_n < L + \epsilon_0$$

for all integers $n = N, N+1, N+2, N+3, \dots$

Thus to find a number M larger than all the values of $|s_n|$ we can select the maximum of these numbers:

$$|s_1|, |s_2|, |s_3|, \dots, |s_{N-2}|, |s_{N-1}|, |L| + \epsilon_0.$$

The simplest bounded sequence that is not convergent would be $s_n = (-1)^n$. It is clearly bounded and obviously violates the definition of convergent.

Exercise 15, page 6

If a sequence of real numbers $\{s_n\}$ is Cauchy then, for any choice of $\epsilon_0 > 0$ there is an integer N so that

$$|s_n - s_N| < \epsilon_0$$

for all integers $n = N, N+1, N+2, N+3, \dots$

Thus to find a number M larger than all the values of $|s_n|$ we can select the maximum of these numbers:

$$|s_1|, |s_2|, |s_3|, \dots, |s_{N-2}|, |s_{N-1}|, |s_N| + \epsilon_0.$$

The simplest bounded sequence that is not Cauchy would be $s_n = (-1)^n$. It is clearly bounded and obviously violates the definition of a Cauchy sequence.

Exercise 16, page 6

If a sequence of real numbers $\{s_n\}$ converges to a real number L then, for every $\epsilon > 0$ there is an integer N so that

$$L - \epsilon/2 < s_n < L + \epsilon/2$$

for all integers $n \geq N$.

Now consider pairs of integers $n, m \geq N$. We compute that

$$|s_n - s_m| = |s_n - L + L - s_m| \leq |s_n - L| + |L - s_m| < \epsilon.$$

By definition then $\{s_n\}$ is a Cauchy sequence.

Exercise 17, page 6

The easiest of these is the formula

$$\lim_{n \rightarrow \infty} (as_n + bt_n) = a \left(\lim_{n \rightarrow \infty} s_n \right) + b \left(\lim_{n \rightarrow \infty} t_n \right).$$

You should certainly review your studies of sequence limits if it does not immediately occur to you how to prove this using the definition of limit.

The product $s_n t_n$ and quotient $\frac{s_n}{t_n}$ formulas are a little harder to prove and require a bit of thinking about the inequalities. Make that when you state and try to prove the quotient formula

$$\lim_{n \rightarrow \infty} \frac{s_n}{t_n} = \frac{\lim_{n \rightarrow \infty} s_n}{\lim_{n \rightarrow \infty} t_n}$$

you include an hypothesis to exclude division by zero on either side of the identity.

Exercise 18, page 7

We already know by an earlier exercise that a convergent sequence would have to be bounded, so it is enough for us to prove that on the assumption that this sequence is bounded it must converge.

Since the sequence is bounded

$$L = \sup\{s_n : n = 1, 2, 3, \dots\}$$

is a real number. It has the property (as do all suprema) that $s_n \leq L$ for all n and, if $\varepsilon > 0$, then $s_n > L - \varepsilon$ for some n .

Choose any integer N such that $s_N > L - \varepsilon$. Then for all integers $n \geq N$,

$$L - \varepsilon < s_N \leq s_n \leq L < L + \varepsilon.$$

By definition, then,

$$\lim_{n \rightarrow \infty} s_n = L.$$

Notice that if the sequence is unbounded then

$$\sup\{s_n : n = 1, 2, 3, \dots\} = \infty$$

and the sequence diverges to ∞ ,

$$\lim_{n \rightarrow \infty} s_n = \infty.$$

You should be able to give a precise proof of this that refers to Definition 1.3.

Exercise 19, page 7

The condition on the intervals immediately shows that the two sequences $\{a_n\}$ and $\{b_n\}$ are bounded and monotone. The sequence $\{a_n\}$ is monotone nondecreasing and bounded above by b_1 ; the sequence $\{b_n\}$ is monotone nonincreasing and bounded below by a_1 .

By Exercise 18 these sequences converge. Take either $z = \lim_{n \rightarrow \infty} a_n$ or $z = \lim_{n \rightarrow \infty} b_n$. This point is in all of the intervals. The assumption that

$$\lim_{n \rightarrow \infty} (b_n - a_n) = 0$$

makes it clear that only one point can be in all of the intervals.

Exercise 20, page 7

We construct first a nonincreasing subsequence if possible. We call the m th element s_m of the sequence $\{s_n\}$ a turn-back point if all later elements are less than or equal to it, in symbols if $s_m \geq s_n$ for all $n > m$. If there is an infinite subsequence of turn-back points $s_{m_1}, s_{m_2}, s_{m_3}, s_{m_4}, \dots$ then we have found our nonincreasing subsequence since

$$s_{m_1} \geq s_{m_2} \geq s_{m_3} \geq s_{m_4} \geq \dots$$

This would not be possible if there are only finitely many turn-back points. Let us suppose that s_M is the last turn-back point so that any element s_n for $n > M$ is not a turn-back point. Since it is not there must be an element further on in the sequence greater than it, in symbols $s_m > s_n$ for some $m > n$. Thus we can choose $s_{m_1} > s_{M+1}$ with $m_1 > M + 1$, then $s_{m_2} > s_{m_1}$ with $m_2 > m_1$, and then $s_{m_3} > s_{m_2}$ with $m_3 > m_2$, and so on to obtain an increasing subsequence

$$s_{M+1} < s_{m_1} < s_{m_2} < s_{m_3} < s_{m_4} < \dots$$

as required.

Exercise 21, page 7

This follows immediately from Exercises 18 and 20. Take any monotone subsequence. Any one of them converges by Exercise 18 since the sequence and the subsequence must be bounded.

Exercise 22, page 7

By the way, before seeing a hint you might want to ask for a reason for the terminology. If every Cauchy sequence is convergent and every convergent sequence is Cauchy why bother with two words for the same idea. The answer is that this same language is used in other parts of mathematics where every convergent sequence is Cauchy, but not every Cauchy sequence is convergent. Since we are on the real line in this course we don't have to worry about such unhappy possibilities. But we retain the language anyway.

What is most important for you to remember is the logic of this exercise so we will sketch that and leave the details for you to write out:

1. Every Cauchy sequence is bounded.
2. Every sequence has a monotone subsequence.
3. Every bounded, monotone sequence converges.
4. Therefore every Cauchy sequence has a convergent subsequence.
5. When a Cauchy sequence has a subsequence converging to a number L the sequence itself must converge to the number L . [Using an ϵ, N argument.]

Exercise 23, page 8

If x does not belong to E then it belongs to a component interval (a, b) of $\mathbb{R} \setminus E$ that contains no points of E . Thus there is a $\delta > 0$ so that $(x - \delta, x + \delta)$ does not contain any points of E . Since all points in the sequence $\{x_n\}$ belong to E this would contradict the statement that $x = \lim_{n \rightarrow \infty} x_n$.

Exercise 24, page 9

Just notice that

$$S_n - S_m = \sum_{k=m-1}^n a_k$$

provided that $n \geq m$.

Exercise 25, page 9

First observe, by the triangle inequality that

$$\left| \sum_{k=m}^n a_k \right| \leq \sum_{k=m}^n |a_k|.$$

Then if we can choose an integer N so that

$$\sum_{k=m}^n |a_k| < \varepsilon$$

for all $n \geq m \geq N$, we can deduce immediately that

$$\left| \sum_{k=m}^n a_k \right| < \varepsilon$$

for all $n \geq m \geq N$.

What can we conclude? If the series of absolute values

$$\sum_{k=1}^n |a_k| = |a_1| + |a_2| + |a_3| + |a_4| + \dots$$

converges then it must follow, without further checking, that the original series

$$\sum_{k=1}^n a_k = a_1 + a_2 + a_3 + a_4 + \dots$$

is also convergent. Thus to determine whether a series $\sum_{k=1}^n a_k$ is absolutely convergent we need only check the corresponding series of absolute values.

Exercise 26, page 10

Just choose any finite number of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

so that the points are closer together than δ . Then no matter what points ξ_i in $[x_{i-1}, x_i]$ we choose the partition

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of the interval $[a, b]$ has the property that each interval $[x_{i-1}, x_i]$ has length smaller than $\delta(\xi_i) = \delta$.

Note that this construction reveals just how hard it might seem to arrange for a partition if the values of $\delta(x)$ are allowed to vary.

Exercise 27, page 10

For every point x in a closed, bounded interval $[a, b]$ let there be given a positive number $\delta(x)$. Let us call an interval $[c, d] \subset [a, b]$ a *black interval* if there exists at least one partition

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of the interval $[c, d]$ with the property that each interval $[x_{i-1}, x_i]$ has length smaller than $\delta(\xi_i)$. If an interval is not black let us say it is *white*.

Observe these facts about black intervals.

1. If $[c, d]$ and $[d, e]$ are black then $[c, e]$ is black.
2. If $[c, d]$ contains a point z for which $d - c < \delta(z)$ then $[c, d]$ is black.

The first statement follows from the fact that any partitions for $[c, d]$ and $[d, e]$ can be joined together to form a partition of $[c, e]$. The second statement follows from the fact that $\{[c, d], z\}$ alone makes up a partition satisfying the required condition in the Cousin lemma.

Now here is the nested interval argument. We wish to prove that $[a, b]$ is black. If it is not black then one of the two intervals $[a, \frac{1}{2}(a + b)]$ or $[\frac{1}{2}(a + b), b]$ is white. If both were black then statement (1) makes $[a, b]$ black. Choose that interval (the white one) as $[a_1, b_1]$. Divide that interval into half again and produce another white interval of half the length. This produces $[a_1, b_1] \supset [a_2, b_2] \supset [a_3, b_3] \dots$, a shrinking sequence of white intervals with lengths decreasing to

zero,

$$\lim_{n \rightarrow \infty} (b_n - a_n) = 0.$$

By the nested interval argument then there is a unique point z that belongs to each of the intervals and there must be an integer N so that $(b_N - a_N) < \delta(z)$. By statement (2) above that makes $[a_N, b_N]$ black which is a contradiction.

Exercise 28, page 10

For every point x in a closed, bounded interval $[a, b]$ let there be given a positive number $\delta(x)$. Let us say that a number $a < r \leq b$ can be reached if there exists at least one partition

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of the interval $[a, r]$ with the property that each interval $[x_{i-1}, x_i]$ has length smaller than $\delta(\xi_i)$.

Define R as the last point that can be reached, i.e.,

$$R = \sup\{r : a < r \leq b \text{ and } r \text{ can be reached}\}.$$

This set is not empty since all points in $(a, a + \delta(a))$ can be reached. Thus R is a real number no larger than b . Check that R itself can be reached. Indeed there must be points r in the interval $(R - \delta(R), R]$ that can be reached (by the definition of sups). If $R - \delta(R) < r < R$ and r can be reached, then R also can be reached by simply adding the element $([r, R], R)$ to the partition for $[a, r]$.

Is $R < b$? No since if it were then we could reach a bigger point by adding a suitable pair $([R, s], R)$ to a partition for $[a, R]$. Consequently $R = b$ and b can be reached, i.e., it is the last point that can be reached.

Exercise 29, page 10

For each x in $[a, b]$ select a positive number $\delta(x)$ so that the open interval

$$(x - \delta(x), x + \delta(x))$$

is inside some open interval of the family $\mathcal{C}$.

By Cousin's lemma there exists at least one partition

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of the interval $[a, b]$ with the property that each interval $[x_{i-1}, x_i]$ has length smaller than $\delta(\xi_i)$. For each $i = 1, 2, 3, \dots$ select from $\mathcal{C}$ some open interval (c_i, d_i) that contains

$$(\xi_i - \delta(\xi_i), \xi_i + \delta(\xi_i)).$$

This finite list of intervals from $\mathcal{C}$

$$(c_1, d_1), (c_2, d_2), \dots, (c_n, d_n)$$

contains every point of $[a, b]$ since every interval $[x_{i-1}, x_i]$ is contained in one of these open intervals.

Exercise 30, page 10

Use a proof by contradiction. For example if $(a, b) \subset G_1 \cup G_2$, $G_1 \cup G_2 = \emptyset$, then for every $x \in G_1 \cap [a, b]$ there is a $\delta(x) > 0$ so that

$$(x - \delta(x), x + \delta(x)) \subset G_1$$

and for every $x \in G_2 \cap [a, b]$ there is a $\delta(x) > 0$ so that

$$(x - \delta(x), x + \delta(x)) \subset G_2.$$

By Cousin's lemma there exists at least one partition

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of the interval $[a, b]$ with the property that each interval $[x_{i-1}, x_i]$ has length smaller than $\delta(\xi_i)$. Consequently each interval $[x_{i-1}, x_i]$ belongs entirely either to G_1 or belongs entirely either to G_2 .

This is impossible. For if $a \in G_1$, then $[x_0, x_1] \subset G_1$. But that means $[x_1, x_2] \subset G_1$, and $[x_2, x_3] \subset G_1, \dots$, and indeed all of the intervals are subsets of G_1 .

Exercise 31, page 11

Use a proof by contradiction. Suppose, for example that $[a, b] \subset G_1 \cup G_2$, $G_1 \cup G_2 = \emptyset$. Then a is in one of these two open sets, say $a \in G_1$. Take the last point t for which $[a, t] \subset G_1$, i.e.,

$$t = \sup\{r : a < r \leq b, [a, r] \subset G_1\}.$$

That number cannot be b , otherwise G_2 contains no point of the interval. And that number cannot be in the open set G_1 , otherwise we failed to pick the last such number. Thus $t \in G_2$. But the situation $t \in G_2$ requires there to be some interval (c, d) containing t and entirely contained inside G_2 . That gives us $(c, t) \subset G_1$ and $(c, t) \subset G_2$. This is a contradiction to the requirement that $G_1 \cup G_2 = \emptyset$.

Exercise 32, page 11

Take any two points in the set, $s < t$. If there is a point $s < c < t$ that is not in the set E then $E \subset (-\infty, c) \cup (c, \infty)$ exhibits, by definition, that the set is disconnected. So the set E contains all points between any two of its elements. Consequently E is either (a, b) or $[a, b)$ or $(a, b]$ or $[a, b]$ where for a take $\inf E$ and for b take $\sup E$.

Exercise 33, page 12

You should remember that these functions are defined for all real numbers, with the exception that $\tan(\pm\pi/2) = \pm\infty$. So, since we do not consider functions to have infinite values, the function $\tan x$ is considered to be defined at all reals that are not of the form $(n + 1/2)\pi/2$ for some integer n .

Exercise 34, page 12

The value of $\arcsin x$ is defined to be the number $-\pi/2 \leq y \leq \pi/2$ such that $\sin y = x$. The only numbers x that permit a solution to this equation are from the interval $[-1, 1]$. The value of $\arctan x$ is defined to be the number $-\pi/2 \leq y \leq \pi/2$ such that $\tan y = x$. This equation can be solved for all real numbers x so the assumed domain of $\arctan x$ is the entire real line.

Exercise 35, page 12

The exponential function e^x is defined for all values of x so its domain is the whole real line. The logarithm function is the inverse defined by requiring $\log x = y$ to mean $e^y = x$. Since e^y is always positive the logarithm function cannot be defined at zero or at any negative number. In fact the domain of $\log x$ is the open unbounded interval $(0, \infty)$.

Exercise 36, page 12

Check that $x^2 - x - 1$ only at the points $c_1 = 1/2 + \sqrt{5}/2$ and $c_2 = 1/2 - \sqrt{5}/2$. Thus the domain of the first function would be assumed to be $(-\infty, c_2) \cup (c_2, c_1) \cup (c_1, \infty)$.

Also $x^2 - x - 1 \geq 0$ only on the intervals $(\infty, c_2]$ and $[c_1, \infty)$ so the the domain of the second function would be assumed to be $(\infty, c_2] \cup [c_1, \infty)$.

Finally, the third function is a composition. We cannot write $\arcsin t$ unless $-1 \leq t \leq 1$, consequently we cannot write $\arcsin(x^2 - x - 1)$ unless $-1 \leq x^2 - x - 1 \leq 1$, or equivalently $0 \leq x^2 - x \leq 2$.

But $-1 \leq x^2 - x - 1$ on the intervals $(-\infty, 0)$ and $(1, \infty)$, while $x^2 - x - 1 \leq 1$ on the interval $[-1, 2]$. So finally $\arcsin(x^2 - x - 1)$ can only be written for x in the intervals $[-1, 0]$ and $[1, 2]$. So the the domain of the third function would be assumed to be $[-1, 0] \cup [1, 2]$.

Exercise 37, page 14

This is trivial. Just state the definition of uniform continuity and notice that it applies immediately to every point.

Exercise 38, page 14

Find a counterexample, i.e., find a function that is continuous on some open interval I and that is not necessarily uniformly continuous on that interval.

Exercise 39, page 14

There are lots of choices. Our favorite might be to set $f(x) = x$ if x is rational and $f(x) = -x$ if x is irrational. Then just check the definition at each point.

Exercise 40, page 14

To work with the definition one must know it precisely and also have an intuitive grasp. Usually we think that uniform continuity of f means

...if $d - c$ is small enough $f(d) - f(c)$ should be small.

For the function $f(x) = x$ this becomes

... if $d - c$ is small enough $d - c$ should be small.

That alone is enough to indicate that the exercise must be trivial. Just write out the definition using $\delta = \epsilon$.

Exercise 41, page 14

Obtain a contradiction by assuming [falsely] that $f(x) = x^2$ is uniformly continuous on the interval $(-\infty, \infty)$.

Usually we think that uniform continuity of f means

... if $d - c$ is small enough $f(d) - f(c)$ should be small.

That means that the failure of uniform continuity should be thought of this way:

... even though $d - c$ is small $f(d) - f(c)$ might not be small.

For the function $f(x) = x^2$ this becomes

... even though $d - c$ is small $d^2 - c^2$ might not be small.

A similar way of thinking is

... even though t is small $(x+t)^2 - x^2$ might not be small.

That should be enough to indicate a method of answering the exercise.

Thus, take any particular $\epsilon > 0$ and suppose [wrongly] that

$$|d^2 - c^2| < \epsilon$$

whenever c, d are points for which $|d - c| < \delta$. Take any large integer N so that $1/N < \delta$. Then

$$|(N + 1/N)^2 - 1/N^2| = N^2 + 2 < \epsilon.$$

This cannot be true for all large integers N so we have a contradiction.

By the way, this method of finding two sequences $x_N = N$ and $y_N = N + 1/N$ to show that uniform continuity fails is turned into a general method in Exercise 90.

Exercise 42, page 14

The key is to factor

$$x^2 - y^2 = (x + y)(x - y).$$

Then, we think that uniform continuity of f means

...if $x - y$ is small enough $f(x) - f(y)$ should be small.

For the function $f(x) = x^2$ this becomes

...if $x - y$ is small enough $[x + y](x - y)$ should be small.

In any bounded interval we can control the size of $[x + y]$.

Here is a formal proof using this thinking. Let I be a bounded interval and suppose that $|x| \leq M$ for all $x \in I$. Let $\varepsilon > 0$ and choose $\delta = \varepsilon/(2M)$. Then, if $|d - c| < \delta$

$$|f(d) - f(c)| = |d^2 - c^2| = |[d + c](d - c)| \leq [|d| + |c|]|d - c| \leq [2M]|d - c| \leq 2M\delta = \varepsilon.$$

By definition, f is uniformly continuous on I .

Exercise 43, page 14

We have already proved that this function is uniformly continuous on any bounded interval. Use that fact on the interval $(x_0 - 1, x_0 + 1)$.

Exercise 44, page 14

Suppose [falsely] that $f(x) = \frac{1}{x}$ is uniformly continuous on the interval $(0, \infty)$; then it must also be uniformly continuous on the bounded interval $(0, 1)$. Using $\varepsilon = 1$ choose $\delta > 0$ so that

$$\left| \frac{1}{x} - \frac{1}{y} \right| < 1$$

if $|x - y| < \delta$. In particular take a point $0 < y_0 < \delta$ and notice that

$$\left| \frac{1}{x} \right| < 1 + \frac{1}{y_0}$$

for all $0 < x < \delta$. For $\delta \leq x < 1$

$$\left| \frac{1}{x} \right| \leq \frac{1}{\delta}$$

We know that this function $f(x) = \frac{1}{x}$ is unbounded and yet we seem to have produced an upper bound on the interval $(0, 1)$. This is a contradiction and hence the function cannot be uniformly continuous.

In fact we can make this particular observation into a method. When a function is uniformly continuous on a bounded interval we will prove that the function is bounded. Hence unbounded functions cannot be uniformly continuous on a bounded interval.

We now show that $f(x) = \frac{1}{x}$ is continuous at every real number $x_0 \neq 0$. Take any point $x_0 > 0$ and let $\epsilon > 0$. We must choose a $\delta > 0$ so that

$$|1/x - 1/x_0| < \epsilon$$

whenever x is a point in $(0, \infty)$ for which $|x - x_0| < \delta$. This an exercise in inequalities. Write

$$|1/x - 1/x_0| = \left| \frac{x - x_0}{xx_0} \right|.$$

Note that if $x > x_0/2$ then $xx_0/2 > x_0^2$ so that $1/[xx_0] \leq 1/[2x_0^2]$. These inequalities reveal the correct choice of δ and reveal where we should place the argument. We need not work in the entire interval $(0, \infty)$ but can restrict the argument to the subinterval $(x_0/2, 3x_0/2)$.

Let x_0 be a point in the interval $(0, \infty)$. Work entirely inside the interval $(x_0/2, 3x_0/2)$. Let $\epsilon > 0$. Choose $\delta = \epsilon x_0^2$ and suppose that $|x - x_0| < \delta = \epsilon x_0^2$ and that x is a point in the interval $(x_0/2, 3x_0/2)$. Then since $x_0/2 < x$,

$$|f(x) - f(x_0)| = \left| \frac{x - x_0}{xx_0} \right| \leq \frac{\epsilon x_0^2}{2x_0^2} < \epsilon.$$

By definition $f(x) = \frac{1}{x}$ is continuous at the point x_0 .

Note this device used here: since pointwise continuity at x_0 is a local property at a point we can restrict the argument to *any* open interval that contains x_0 . If, by doing so, you can make the inequality work easier then, certainly, do so.

Note: We have gone into some great detail in the exercise since this is at an early stage in our theory and it is an opportunity for instruction. You should be able to write up this proof in a shorter, more compelling presentation.

Exercise 45, page 14

It is easy to check that both $rF(x)$ and $sG(x)$ must be continuous at the point x_0 . Thus it is enough to prove the result for $r = s = 1$, i.e., to prove that $F(x) + G(x)$ must be continuous at the point x_0 .

The inequality

$$|F(x) + G(x) - [F(x_0) + G(x_0)]| \leq |F(x) - F(x_0)| + |G(x) - G(x_0)|$$

suggests an easy proof.

Using the same method you should be successful in proving the following statement:

Let F and G be functions that are uniformly continuous on an interval I . Then any linear combination $H(x) = rF(x) + sG(x)$ must also be uniformly continuous on I .

Exercise 46, page 14

The key is to use the simple inequality

$$\begin{aligned} |F(x)G(x) - F(x_0)G(x_0)| &= |F(x)G(x) - F(x_0)G(x) + F(x_0)G(x) - F(x_0)G(x_0)| \\ &\leq |F(x)G(x) - F(x_0)G(x)| + |F(x_0)G(x) - F(x_0)G(x_0)| \\ &= |G(x)| \cdot |F(x) - F(x_0)| + |F(x_0)| \cdot |G(x) - G(x_0)| \end{aligned}$$

Since G is continuous at the point x_0 there must be at least one interval $(c, d) \subset I$ containing the point x_0 so that G is bounded on (c, d) . In fact we can use the definition of continuity to find a $\eta > 0$ so that

$$|G(x) - G(x_0)| < 1 \text{ for all } x \text{ in } (x_0 - \eta, x_0 + \eta)$$

and so, also

$$|G(x)| < |G(x_0)| + 1 \text{ for all } x \text{ in } (x_0 - \eta, x_0 + \eta).$$

Thus we can select such an interval (c, d) and a positive number M that is larger than $|G(x)| + |F(x_0)|$ for all x in the interval (c, d) .

Let $\varepsilon > 0$. The assumptions imply the existence of the positive numbers δ_1 and δ_2 , such that

$$|F(x) - F(x_0)| < \frac{\varepsilon}{2M}$$

if $|x - x_0| < \delta_1$ and

$$|G(x) - G(x_0)| < \frac{\varepsilon}{2M}$$

if $|x - x_0| < \delta_2$.

Then, using any δ smaller than both δ_1 and δ_2 , and arguing inside the interval (c, d) we observe that

$$|F(x)G(x) - F(x_0)G(x_0)| \leq M|F(x) - F(x_0)| + M|G(x) - G(x_0)| < 2M\varepsilon/(2M) = \varepsilon.$$

if $|x - x_0| < \delta$. This is immediate from the inequalities above. This proves that the product $H(x) = F(x)G(x)$ must be continuous at the point x_0 .

Does the same statement apply to uniform continuity? In view of Exercise 45 you might be tempted to prove the following false theorem:

FALSE: Let F and G be functions that are uniformly continuous on an interval I . Then the product $H(x) = F(x)G(x)$ must also be uniformly continuous on I .

But note that $F(x) = G(x) = x$ are both uniformly continuous on $(-\infty, \infty)$ while $FG(x) = F(x)G(x) = x^2$ is not. The key is contained in your proof of this exercise. You needed boundedness to make the inequalities work.

Here is a true version that you can prove using the methods that we used for the pointwise case:

TRUE: Let F and G be functions that are uniformly continuous on an interval I . Suppose that G is bounded on the interval I . Then the product $H(x) = F(x)G(x)$ must also be uniformly continuous on I .

Later on we will find that, when working on bounded intervals, all uniformly continuous functions must be bounded. If you use this fact now and repeat your arguments you can prove the following version:

Let F and G be functions that are uniformly continuous on a bounded interval I . Then the product $H(x) = F(x)G(x)$ must also be uniformly continuous on I .

Exercise 47, page 14

Yes, if $G(x_0) \neq 0$. The identity

$$\left| \frac{F(x)}{G(x)} - \frac{F(x_0)}{G(x_0)} \right| = \left| \frac{F(x)G(x_0) - F(x_0)G(x) + F(x)G(x) - F(x_0)G(x)}{G(x)G(x_0)} \right|.$$

should help. You can also prove the following version for uniform continuity.

Let F and G be functions that are uniformly continuous on an interval I . Then the quotient $H(x) = F(x)/G(x)$ must also be uniformly continuous on I provided that the functions F and $1/G$ are also defined and bounded on I .

Exercise 48, page 14

Let $\varepsilon > 0$ and determine $\eta > 0$ so that

$$|G(z) - G(z_0)| < \varepsilon$$

whenever z is a point in J for which $|z - z_0| < \eta$. Now use the continuity of F at the point x_0 to determine a $\delta > 0$ so that

$$|F(x) - F(x_0)| < \eta$$

whenever x is a point in I for which $|x - x_0| < \delta$.

Note that if x is a point in I for which $|x - x_0| < \delta$, then $z = F(x)$ is a point in J for which $|z - z_0| < \eta$. Thus

$$|G(F(x)) - G(F(x_0))| = |G(z) - G(z_0)| < \varepsilon.$$

Exercise 51, page 15

Too simple for a hint.

Exercise 52, page 15

Another way to think about this is that a function that is a sum of characteristic functions

$$f(x) = \sum_{i=1}^M a_i \chi_{A_i}(x)$$

is a step function if all the A_i are intervals or singleton sets. [Here $\chi_E(x)$, for a set E , is equal to 1 for points x in E and equal to 0 otherwise.]

Let $f : [a, b] \rightarrow \mathbb{R}$ be a step function. Show first that there is a partition

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

so that f is constant on each interval (x_{i-1}, x_i) , $i = 1, 2, \dots, n$. This will display all possible discontinuities.

Exercise 53, page 16

This is not so hard and the title gives it away. Show first that $R(x) = 1$ if x is rational and is otherwise 0.

Exercise 54, page 16

We can interpret this statement, that the distance function is continuous, geometrically this way: if two points x_1 and x_2 are close together, then they are at roughly the same distance from the closed set C .

Exercise 55, page 16

This just requires connecting two definitions: the definition of continuity at a point and the definition of sequential limit at a point.

Let $\varepsilon > 0$ and choose $\delta > 0$ so that $|f(x) - f(x_0)| < \varepsilon$ if $|x - x_0| < \delta$. Now choose N so that $|x_n - x_0| < \delta$ for $n > N$. Combining the two we get that $|f(x_n) - f(x_0)| < \varepsilon$ if $n > N$. By definition that means that $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$.

That proves one direction. To prove the other direction we can use a contrapositive argument: assume that continuity fails and then deduce that the sequence property also fails. Suppose that f is not continuous at x_0 . Then, for some value of ε there cannot be a δ for which $|f(x) - f(x_0)| < \varepsilon$ if $|x - x_0| < \delta$. Consequently, for every integer n there must be at least one point x_n in the interval so that $|x_0 - x_n| < 1/n$ and yet $|f(x) - f(x_0)| > \varepsilon$.

In other words we have produced a sequence $\{x_n\} \rightarrow x_0$ for which $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$ fails.

Exercise 56, page 16

Suppose first that f is continuous. Let V be open, let $x_0 \in f^{-1}(V)$ and choose $\alpha < \beta$ so that $(\alpha, \beta) \subset V$ and so that $x_0 \in f^{-1}((\alpha, \beta))$. Then $\alpha < f(x_0) < \beta$. We will find a neighborhood U of x_0 such that $\alpha < f(x) < \beta$ for all $x \in U$. Let

$$\varepsilon = \min(\beta - f(x_0), f(x_0) - \alpha).$$

11.2. ANSWERS TO EXERCISES

Since f is continuous at x_0 , there exists a $\delta > 0$ such that if

$$x \in (x_0 - \delta, x_0 + \delta),$$

then

$$|f(x) - f(x_0)| < \varepsilon.$$

Thus

$$f(x) - f(x_0) < \beta - f(x_0),$$

and so $f(x) < \beta$. Similarly,

$$f(x) - f(x_0) > \alpha - f(x_0),$$

and so $f(x) > \alpha$. Thus the neighborhood $U = (x_0 - \delta, x_0 + \delta)$ is a subset of $f^{-1}((\alpha, \beta))$ and hence also a subset of $f^{-1}(V)$. We have shown that each member of $f^{-1}(V)$ has a neighborhood in $f^{-1}(V)$. That is, $f^{-1}(V)$ is open.

To prove the converse, suppose f satisfies the condition that for each open interval (α, β) with $\alpha < \beta$, the set $f^{-1}((\alpha, \beta))$ is open. Take a point x_0 . We must show that f is continuous at x_0 . Let $\varepsilon > 0$, $\beta = f(x_0) + \varepsilon$, $\alpha = f(x_0) - \varepsilon$. Our hypothesis implies that $f^{-1}((\alpha, \beta))$ is open.

Thus there is at least open interval, (c, d) say, that is contained in this open set and contains the point x_0 . Let

$$\delta = \min(x_0 - c, d - x_0).$$

For $|x - x_0| < \delta$ we find

$$\alpha < f(x) < \beta.$$

Because $\beta = f(x_0) + \varepsilon$ and $\alpha = f(x_0) - \varepsilon$ we must have

$$|f(x) - f(x_0)| < \varepsilon.$$

This shows that f is continuous at x_0 .

Exercise 58, page 17

In preparation ...

Exercise 59, page 17

Because of Exercise 58 we already know that if $\varepsilon > 0$ then there is $\delta > 0$ so that

$$\omega f([c, d]) < \varepsilon$$

whenever $[c, d]$ is a subinterval of I for which $|d - c| < \delta$. If the points of the subdivision

$$a = x_0 < x_1 < x_2 < x_3 < \cdots < x_{n-1} < x_n = b$$

are chosen with gaps smaller than δ then, certainly, each of

$$\omega f([x_0, x_1]), \omega f([x_1, x_2]), \dots, \text{ and } \omega f([x_{n-1}, x_n])$$

is smaller than ε .

Conversely suppose that there is such a subdivision. Let δ be one-half of the minimum of the lengths of the intervals $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$. Note that if we take any interval $[c, d]$ with length less than δ that interval can meet no more than two of the intervals above. For example if $[c, d]$ meets both $[x_0, x_1]$ and $[x_1, x_2]$, then

$$\omega f([c, d]) \leq \omega f([x_0, x_1]) + \omega f([x_1, x_2]) < 2\varepsilon.$$

In fact then any interval $[c, d]$ with length less than δ must have

$$\omega f([c, d]) < 2\varepsilon.$$

It follows that f is uniformly continuous on $[a, b]$.

Is there a similar statement for uniform continuity on open intervals? Yes. Just check that f is a uniformly continuous function on an open, bounded interval (a, b) if and only if, for every $\varepsilon > 0$, there are points

$$a = x_0 < x_1 < x_2 < x_3 < \cdots < x_{n-1} < x_n = b$$

so that each of

$$\omega f([x_0, x_1]), \omega f([x_1, x_2]), \dots, \text{ and } \omega f([x_{n-1}, x_n])$$

is smaller than ε .

Exercise 60, page 18

If f is continuous at a point x_0 and $\varepsilon > 0$ there is a $\delta(x_0) > 0$ so that

$$|f(x) - f(x_0)| < \varepsilon/2$$

for all $|x - x_0| \leq \delta(x_0)$. Take any two points u and v in the interval $[x_0 - \delta(x_0), x_0 + \delta(x_0)]$ and check that

$$|f(v) - f(u)| \leq |f(v) - f(x_0)| + |f(v) - f(x_0)| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

It follows that

$$\omega f([x_0 - \delta(x_0), x_0 + \delta(x_0)]) \leq \varepsilon.$$

The other direction is easier since

$$|f(x) - f(x_0)| \leq \omega f([x_0 - \delta(x_0), x_0 + \delta(x_0)])$$

for all $|x - x_0| \leq \delta(x_0)$.

Exercise 61, page 18

This is just a rephrasing of the previous exercise.

Exercise 62, page 18

We use the fact that one-sided limits and sequential limits are equivalent in this sense:

A necessary and sufficient condition in order that

$$L = \lim_{x \rightarrow a+} F(x)$$

should exist is that for all decreasing sequence of points $\{x_n\}$ convergent to a , the sequence $\{f(x_n)\}$ converges to L .

Let us prove the easy direction first. Suppose that $F(a+) = \lim_{x \rightarrow a+} F(x)$ exists and let $\varepsilon > 0$. Choose $\delta(a) > 0$ so that

$$|F(a+) - F(x)| \leq \varepsilon/3$$

for all $a < x < a + \delta(a)$. Then, for all $c, d \in (a, a + \delta(a))$,

$$|F(d) - F(c)| \leq |F(a+) - F(d)| + |F(a+) - F(c)| \leq 2\varepsilon/3.$$

It follows that

$$\omega F((a, a + \delta(a))) \leq 2\varepsilon/3 < \varepsilon.$$

In the other direction consider a decreasing sequence of points $\{x_n\}$ convergent to a . Let $\varepsilon > 0$ and choose $\delta(a)$ so that

$$\omega F((a, a + \delta(a))) < \varepsilon.$$

Then there is an integer N so that $|x_n - a| < \delta(a)$ for all $n \geq N$. Thus

$$|F(x_n) - F(x_m)| \leq \omega F((a, a + \delta(a))) < \varepsilon$$

for all $m, n \geq N$. It follows from the Cauchy criterion for sequences that every such sequence $\{F(x_n)\}$ converges. The limit is evidently $f(a+)$.

Exercise 63, page 18

We use the fact that infinite limits and sequential limits are equivalent in this sense:

A necessary and sufficient condition in order that

$$L = \lim_{x \rightarrow \infty} F(x)$$

should exist is that for all sequence of points $\{x_n\}$ divergent to ∞ , the sequence $\{F(x_n)\}$ converges to L .

Let us prove the easy direction first. Suppose that $F(\infty) = \lim_{x \rightarrow \infty} F(x)$ exists and let $\varepsilon > 0$. Choose $T > 0$ so that

$$|F(\infty) - F(x)| \leq \varepsilon/3$$

for all $T < x$. Then, for all $c, d \in (T, \infty)$,

$$|F(d) - F(c)| \leq |F(\infty) - F(d)| + |F(\infty) - F(c)| \leq 2\varepsilon/3.$$

It follows that

$$\omega F((T, \infty)) \leq 2\varepsilon/3 < \varepsilon.$$

In the other direction consider a sequence of points $\{x_n\}$ divergent to ∞ . Let $\varepsilon > 0$ and choose T so that

$$\omega F((T, \infty)) < \varepsilon.$$

Then there is an integer N so that $x_n > T$ for all $n \geq N$. Thus

$$|F(x_n) - F(x_m)| \leq \omega F((T, \infty)) < \varepsilon$$

for all $m, n \geq N$. It follows from the Cauchy criterion for sequences that every such sequence $\{F(x_n)\}$ converges. The limit is evidently $F(\infty)$.

Exercise 64, page 20

This is a direct consequence of Exercise 62. Let $\varepsilon > 0$ and choose $\delta > 0$ so that

$$\omega F((c, d)) < \varepsilon$$

for all subintervals $[c, d]$ of (a, b) for which $d - c < \delta$.

Then, certainly,

$$\omega F((a, a + \delta)) < \varepsilon \text{ and } \omega F((b - \delta, b)) < \varepsilon.$$

From this, Exercise 62 supplies the existence of the two one-sided limits

$$F(a+) = \lim_{x \rightarrow a+} F(x) \text{ and } F(b-) = \lim_{x \rightarrow b-} F(x).$$

Exercise 65, page 20

Let $\varepsilon > 0$ and, using Exercise 62, choose positive numbers $\delta(a)$ and $\delta(b)$ so that

$$\omega F((a, a + \delta(a))) < \varepsilon \text{ and } \omega F((b - \delta(b), b)) < \varepsilon/2.$$

Now choose, for any point $\xi \in (a, b)$, a positive number and $\delta(\xi)$ so that

$$\omega F([\xi - \delta(\xi), \xi + \delta(\xi)]) < \varepsilon.$$

This just uses the continuity of the function f at the point ξ in the oscillation version of that property that we studied in Section 1.6.2.

By the Cousin partitioning argument there must exist points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

and a partition

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of the whole interval $[a, b]$ such that

$$\xi_i \in [x_{i-1}, x_i] \text{ and } x_i - x_{i-1} < \delta(\xi_i).$$

Just observe that this means that each of the following oscillations is smaller than ε :

$$\omega F((a, x_1]), \omega F([x_1, x_2]), \omega F([x_2, x_3]), \dots, \omega F([x_{n-1}, b)).$$

It follows from Exercise 59 that f is uniformly continuous on (a, b) . In preparation ...

Exercise 66, page 20

In one direction this is trivial. If F is defined on (a, b) but can be extended to a uniformly continuous function on $[a, b]$ then F is already uniformly continuous on (a, b) .

The other direction is supplied by the theorem, in fact in the proof of the theorem. That proof supplied points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

so that each of the following oscillations is smaller than ε :

$$\omega F((a, x_1]), \omega F([x_1, x_2]), \omega F([x_2, x_3]), \dots, \omega F([x_{n-1}, b)).$$

Now define $G = F$ on (a, b) and $G(a) = F(a+)$, $G(b) = F(b-)$. Then

$$\omega F((a, x_1]) = \omega G([a, x_1])$$

and

$$\omega F([x_{n-1}, b)) = \omega G([x_{n-1}, b]).$$

With this rather subtle change we have now produced points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

so that each of the following oscillations is smaller than ε :

$$\omega G([a, x_1]), \omega G([x_1, x_2]), \omega G([x_2, x_3]), \dots, \omega G([x_{n-1}, b]).$$

It follows from Exercise 59 that G is uniformly continuous on $[a, b]$.

Exercise 67, page 20

If F is continuous on (a, b) and $[c, d] \subset (a, b)$ note that F is continuous on $[c, d]$ and that $F(c) = F(c+)$ and $F(d) = F(d-)$. Applying the theorem we see that F is uniformly continuous on $[c, d]$.

Exercise 68, page 20

If F is continuous on (a, b) and monotone nondecreasing then we know that either $F(a+) = \lim_{x \rightarrow a+} F(x)$ exists as a finite real number or else $F(a+) = -\infty$. Similarly know that either $F(b-) = \lim_{x \rightarrow b-} F(x)$ exists as a finite real number or else $F(b-) = +\infty$. Thus, by Theorem 1.10 the function is uniformly continuous on (a, b) provided only that the function is bounded. Conversely, in order for the function F to be uniformly continuous on (a, b) , it must be bounded since all uniformly continuous functions are bounded on bounded intervals.

Exercise 69, page 20

This proof invokes a Bolzano-Weierstrass compactness argument. We use an indirect proof. If F is not uniformly continuous, then there are sequences $\{x_n\}$ and $\{y_n\}$ so that $x_n - y_n \rightarrow 0$ but

$$|F(x_n) - F(y_n)| > c$$

for some positive c . (The verification of this step is left out, but you should supply it. This can be obtained merely by negating the formal statement that f is uniformly continuous on $[a, b]$.)

Now apply the Bolzano-Weierstrass property to obtain a convergent subsequence $\{x_{n_k}\}$. Write z as the limit of this new sequence $\{x_{n_k}\}$. Observe that $x_{n_k} - y_{n_k} \rightarrow 0$ since $x_n - y_n \rightarrow 0$. Thus $\{x_{n_k}\}$ and the corresponding subsequence $\{y_{n_k}\}$ of the sequence $\{y_n\}$ both converge to the same limit z , which must be a point in the interval $[a, b]$.

If $a < z < b$ then we get a contradiction from the continuity of f : $F(x_{n_k}) \rightarrow F(z)$ and $F(y_{n_k}) \rightarrow F(z)$. Since $|F(x_n) -$

$|F(y_n)| > c$ for all n , this means from our study of sequence limits that

$$|F(z) - F(z)| \geq c > 0$$

and this is impossible.

Now suppose that $z = a$. Since

$$F(a+) = \lim_{x \rightarrow a+} F(x)$$

exists it also follows that $F(x_{n_k}) \rightarrow F(a+)$ and $F(y_{n_k}) \rightarrow F(a+)$. Again this is impossible. The remaining case, $z = b$ is similarly handled.

Exercise 70, page 20

Choose open intervals $(a, a + \delta(a))$, $(b - \delta(b), b)$ so that

$$\omega F((a, a + \delta(a))) < \varepsilon/2 \quad \text{and} \quad \omega F((b - \delta(b), b)) < \varepsilon/2$$

At the endpoints a and b this is possible because the one-sided limits exist (i.e., by Exercise 62).

For each point $x \in (a, b)$ we may choose intervals $(x - \delta(x), x + \delta(x))$ in such a way that

$$\omega F((x - \delta(x), x + \delta(x))) < \varepsilon/2.$$

At the points $x \in (a, b)$ this is possible because of our assumption that F is continuous at all such points.

Pick points s and t with $a < s < a + \delta(a)$ and $b - \delta(b) < t < b$. Now apply the Heine-Borel property to this covering of the closed interval $[s, t]$. There are now a finite number of open intervals $(x_i - \delta(x_i), x_i + \delta(x_i))$ with $i = 1, 2, 3, \dots, k$ covering $[s, t]$.

Let δ be half the minimum length of all the intervals

$$(a, a + \delta(a)), (b - \delta(b), b), (x_i - \delta(x_i), x_i + \delta(x_i)) \quad (i = 1, 2, 3, \dots, k).$$

Use this to show that $\omega F([c, d]) < \varepsilon$ if $[c, d] \subset (a, b)$ and $d - c < \delta$.

Exercise 71, page 20

There are functions that are continuous at every point of $(-\infty, \infty)$ and yet are not uniformly continuous. Find one.

Exercise 72, page 20

There are functions that are continuous at every point of $(0, 1)$ and yet are not uniformly continuous. Find one.

Exercise 73, page 21

Slight of hand. Choose δ . Wait a minute. Our choice of δ depended on x_0 so, to make the trick more transparent, call it $\delta(x_0)$. Then if d is some other point you will need a different value of δ .

Exercise 74, page 21

If G is continuous at every point of an interval $[c, d]$ then the theorem (Theorem 1.10) applies to show that G is uniformly continuous on that interval.

Exercise 75, page 21

Just read this from the theorem.

Exercise 76, page 21

Just read this from the theorem.

Exercise 77, page 21

In preparation ...

Exercise 78, page 23

Let f be a uniformly continuous function on a closed, bounded interval $[a, b]$. Take any value of $\epsilon_0 > 0$. Then Exercise 59 supplies points

$$a = x_0 < x_1 < x_2 < x_3 < \cdots < x_{n-1} < x_n = b$$

so that each of

$$\omega f([], \omega f((x_1, x_2]), \dots, \omega f((x_{n-1}, x_n])$$

is smaller than ε_0 . In particular, f is bounded on each of these intervals. Consequently f is bounded on all of $[a, b]$.

The same proof could be used if we had started with a uniformly continuous function on an open bounded interval (a, b) . Note that if the interval is unbounded then such a finite collection of subintervals would not exist.

Exercise 79, page 23

The condition of pointwise continuity at a point x_0 gives us an inequality

$$|f(x) - f(x_0)| < \varepsilon$$

that must hold for some interval $(x_0 - \delta, x_0 + \delta)$. This immediately provides the inequality

$$|f(x)| = |f(x) - f(x_0) + f(x_0)| \leq |f(x) - f(x_0)| + |f(x_0)| < \varepsilon + |f(x_0)|$$

which provides an upper bound for f in the interval $(x_0 - \delta, x_0 + \delta)$.

Exercise 80, page 23

No. It doesn't follow. For a counterexample, the function $f(x) = \sin(1/x)$ is a continuous, bounded function on the bounded open interval $(0, 1)$. This cannot be uniformly continuous because

$$\omega f((0, t)) = 2$$

for every $t > 0$. This function appears again in Exercise 106 The graph of F is shown in Figure 11.1 and helps to illustrate that the continuity cannot be uniform on $(0, 1)$.

Later on we will see that if a function f is uniformly continuous on a bounded open interval (a, b) then the one-sided limits at the endpoints a and b must exist. For the example $f(x) = \sin(1/x)$ on the bounded open interval $(0, 1)$ we can check that $f(0+)$ does not exist, which it must if f were to be uniformly continuous on $(0, 1)$.

Exercise 81, page 23

We assume here that you have studied sequences and convergence of sequences. If f is not bounded then there must be a point x_1 in the interval for which $|f(x_1)| > 1$. If not then $|f(x)| \leq 1$ and we have found an upper bound for the values

of the function. Similarly there must be a point x_2 in the interval for which $|f(x_2)| > 2$, and a point x_3 in the interval for which $|f(x_3)| > 3$. Continue choosing points and then check that $|f(x_n)| \rightarrow \infty$.

Exercise 84, page 23

The function $f(x) = x$ is uniformly continuous on the unbounded interval (∞, ∞) and yet it is not bounded. On the other hand the function $f(x) = \sin x$ is uniformly continuous on the unbounded interval (∞, ∞) and it is bounded, since $|\sin x| \leq 1$.

Exercise 85, page 23

Yes and yes. If $|f(x)| \leq M$ and $|g(x)| \leq N$ for all x in an interval I then

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq M + N$$

and

$$|f(x)g(x)| \leq MN.$$

Exercise 86, page 23

No. On the interval $(0, 1)$ the functions $f(x) = x$ and $g(x) = x^2$ are bounded functions that do not assume the value zero. The quotient function g/f is bounded but the quotient function f/g is unbounded.

Exercise 87, page 23

If the values of $f(t)$ are bounded then the values of $f(g(x))$ are bounded since they include only the same values. Thus there is no need for the extra hypothesis that g is bounded.

Exercise 88, page 24

Use the mean-value theorem to assist in this. If $c < d$ then

$$\sin d - \sin c = (d - c) \cos \xi$$

for some point ξ between c and d .

Don't remember the mean-value theorem? Well use these basic facts instead:

$$\sin d - \sin c = \sin[(d - c)/2] \cos[(d + c)/2]$$

and

$$|\sin x| \leq |x|.$$

Stop reading and try the problem now ...

If you followed either of these hints then you have arrived at an inequality of the form

$$|\sin d - \sin c| \leq M|d - c|.$$

Functions that satisfy this so-called Lipschitz condition are easily shown to be uniformly continuous. For δ you will find that $\delta = \varepsilon/M$ works.

Exercise 89, page 24

For δ you will find that $\delta = \varepsilon/M$ works.

Exercise 90, page 24

If f is uniformly continuous then, by definition, for every $\varepsilon > 0$ there is a $\delta > 0$ so that

$$|f(d) - f(c)| < \varepsilon$$

whenever c, d are points in I for which $|d - c| < \delta$. If there did exist two sequences of points $\{x_n\}$ and $\{y_n\}$ from that interval for which $x_n - y_n \rightarrow 0$ then there would be an integer N so that $|x_n - y_n| < \delta$ for $n > N$. Consequently

$$|f(x_n) - f(y_n)| < \varepsilon$$

for $n > N$. By that means, by the usual sequence definitions that $f(x_n) - f(y_n)$ does indeed converge to zero.

Conversely if f is not uniformly continuous on the interval I then for some value $\varepsilon_0 > 0$ and every integer n the statement that

the inequality

$$|f(d) - f(c)| < \varepsilon_0$$

holds whenever c, d are points in I for which $|d - c| < 1/n$.

must fail. Thus it is possible to select points $\{x_n\}$ and $\{y_n\}$ from that interval for which $|x_n - y_n| < 1/n$ but

$$|f(x_n) - f(y_n)| \geq \varepsilon_0.$$

Consequently we have exhibited two sequences of points $\{x_n\}$ and $\{y_n\}$ from that interval for which $x_n - y_n \rightarrow 0$ but $f(x_n) - f(y_n)$ does not converge to zero.

Exercise 91, page 24

By Theorem 1.16, F is bounded and so we may suppose that M is the least upper bound for the values of F , i.e.,

$$M = \sup\{f(x) : a \leq x \leq b\}.$$

If there exists x_0 such that $F(x_0) = M$, then F achieves a maximum value M . Suppose, then, that $F(x) < M$ for all $x \in [a, b]$. We show this is impossible.

Let $g(x) = 1/(M - F(x))$. For each $x \in [a, b]$, $F(x) \neq M$; as a consequence, g is uniformly continuous and $g(x) > 0$ for all $x \in [a, b]$. From the definition of M we see that

$$\inf\{M - f(x) : x \in [a, b]\} = 0,$$

so

$$\sup\left\{\frac{1}{M - f(x)} : x \in [a, b]\right\} = \infty.$$

This means that g is not bounded on $[a, b]$. This is impossible because uniformly continuous are bounded on bounded intervals. A similar proof would show that F has an absolute minimum.

Exercise 92, page 24

We can also prove Theorem 1.19 using a Bolzano-Weierstrass argument. Let

$$M = \sup\{F(x) : a \leq x \leq b\}.$$

That means that for any integer n the smaller number $M - 1/n$ cannot be an upper bound for the values of the function F on this interval.

Consequently we can choose a sequence of points $\{x_n\}$ from $[a, b]$ so that

$$F(x_n) > M - 1/n.$$

Now apply the Bolzano-Weierstrass theorem to find a subsequence $\{x_{n_k}\}$ that converges to some point z_0 in $[a, b]$. Use the continuity of F to deduce that

$$\lim_{k \rightarrow \infty} F(x_{n_k}) = F(z_0).$$

Since

$$M \geq F(x_{n_k}) > M - 1/n_k$$

it must follow that $F(z_0) = M$. Thus the function F attains its maximum value at z_0 .

Exercise 93, page 24

How about $\sin 2\pi x$? This example is particularly easy to think about since the minimum value could only occur at an endpoint and we have excluded both endpoints by working only on the interval $(0, 1)$. In fact we should notice that this feature is general: if f is a uniformly continuous function on the interval $(0, 1)$ then there is an extension of f to a uniformly continuous function on the interval $[0, 1]$ and the maximum and minimum values are attained on $[0, 1]$ [but not necessarily on $(0, 1)$].

Exercise 94, page 25

How about $1 - \sin 2\pi x$?

Exercise 95, page 25

Simplest is $f(x) = x$.

Exercise 96, page 25

Simplest is $f(x) = x$.

Exercise 97, page 25

Try $f(x) = \arctan x$.

Exercise 98, page 25

If there is a point $f(x_0) = c > 0$, then there is an interval $[-N, N]$ so that $x_0 \in [-N, N]$ and $|f(x)| < c/2$ for all $x > N$ and $x < -N$. Now since f is uniformly continuous on $[-N, N]$ we may select a maximum point. That maximum will be the maximum also on $(-\infty, \infty)$.

If there is no such point x_0 then f assumes only values negative or zero. Apply the same argument but to the function $-f$. For a suitable example of a function that has an absolute maximum but not an absolute minimum you may take $f(x) = (1 + x^2)^{-1}$.

Exercise 99, page 25

All values of $f(x)$ are assumed in the interval $[0, p]$ and f is uniformly continuous on $[0, p]$. It would not be correct to argue that f is uniformly continuous on $(-\infty, \infty)$ [it is] and “hence” that f must have a maximum and minimum [it would not follow].

Exercise 100, page 26

This is a deeper theorem than you might imagine and will require a use of one of our more sophisticated arguments. Try using the Cousin covering argument.

Let f be continuous at points a and b and at all points in between, and let $c \in \mathbb{R}$. If for every $x \in [a, b]$, $f(x) \neq c$, then either $f(x) > c$ for all $x \in [a, b]$ or $f(x) < c$ for all $x \in [a, b]$.

Let $\mathcal{C}$ denote the collection of closed intervals J such that $f(x) < c$ for all $x \in J$ or $f(x) > c$ for all $x \in J$. We verify that $\mathcal{C}$ forms a Cousin cover of $[a, b]$.

If $x \in [a, b]$, then $|f(x) - c| = \varepsilon > 0$, so there exists $\delta > 0$ such that $|f(t) - f(x)| < \varepsilon$ whenever $|t - x| < \delta$ and $t \in [a, b]$. Thus, if $f(x) < c$, then $f(t) < c$ for all $t \in [x - \delta/2, x + \delta/2]$, while if $f(x) > c$, then $f(t) > c$ for all $t \in [x - \delta/2, x + \delta/2]$. By Cousin's lemma there exists a partition of $[a, b]$, $a = x_0 < x_1 < \cdots < x_n = b$ such that for $i = 1, \dots, n$, $[x_{i-1}, x_i] \in \mathcal{C}$.

Suppose now that $f(a) < c$. The argument is similar if $f(a) > c$. Since $[a, x_1] = [x_0, x_1] \in \mathcal{C}$, $f(x) < c$ for all $x \in [x_0, x_1]$. Analogously, since $[x_1, x_2] \in \mathcal{C}$, and $f(x_1) < c$, $f(x) < c$ for $x \in [x_1, x_2]$. Proceeding in this way, we see that $f(x) < c$ for all $x \in [a, b]$.

Exercise 101, page 26

We can prove Theorem 1.21 using the Bolzano-Weierstrass property of sequences rather than Cousin's lemma. Suppose that the theorem is false and explain, then, why there should exist sequences $\{x_n\}$ and $\{y_n\}$ from $[a, b]$ so that $f(x_n) > c$, $f(y_n) < c$ and $|x_n - y_n| < 1/n$.

Exercise 102, page 26

We can prove Theorem 1.21 using the Heine-Borel property. Suppose that the theorem is false and explain, then, why there should exist at each point $x \in [a, b]$ an open interval I_x centered at x so that either $f(t) > c$ for all $t \in I_x \cap [a, b]$ or else $f(t) < c$ for all $t \in I_x \cap [a, b]$.

Exercise 103, page 26

We can prove Theorem 1.21 using the following “last point” argument: suppose that $f(a) < c < f(b)$ and let z be the last point in $[a, b]$ where $f(z)$ stays below c , that is, let

$$z = \sup\{x \in [a, b] : f(t) \leq c \text{ for all } a \leq t \leq x\}.$$

Show that $f(z) = c$.

You may take $c = 0$. Show that if $f(z) > 0$, then there is an interval $[z - \delta, z]$ on which f is positive. Show that if $f(z) < 0$, then there is an interval $[z, z + \delta]$ on which f is negative. Explain why each of these two cases is impossible.

Exercise 104, page 26

For any such function f the Darboux property implies that the image set is connected. In an earlier exercise we determined that all connected sets on the real line are intervals.

For the examples we will need three functions $F, G, H : (0, 1) \rightarrow \mathbb{R}$ so that the image under F is not open, the image under G is not closed, and the image under H is not bounded. You can check that $F(x) = G(x) = x(1 - x)$ maps $(0, 1)$ onto $(0, 1/4]$, and that $H(x) = 1/x$ maps $(0, 1)$ onto $(0, \infty)$.

Exercise 105, page 26

As in the preceding exercise we know that the image set is a connected set [by the Darboux property] and hence that it is an interval. This interval must be bounded since a uniformly continuous function on a closed, bounded interval $[a, b]$ is bounded. This interval must be then either (A, B) or $[A, b)$ or $(A, B]$ or $[A, B]$. The possibilities (A, B) and $(A, B]$ are impossible, for then the function would not have a minimum. The possibilities (A, B) and $[A, B)$ are impossible, for then the function would not have a maximum.

Exercise 106, page 26

The graph of F , shown in Figure 11.1, will help in thinking about this function.

Let us check that F is continuous everywhere, except at $x_0 = 0$. If we are at a point $x_0 \neq 0$ then this function is the composition of two functions $G(x) = \sin x$ and $H(x) = 1/x$, both suitably continuous. So on any interval $[s, t]$ that does not contain $x_0 = 0$ the function is continuous and continuous functions satisfy the Darboux property.

Let $t > 0$. On the interval $[0, t]$, $F(0) = 0$ and F assumes every value between 1 and -1 infinitely often. On the interval $[-t, 0]$, $F(0) = 0$ and F assumes every value between 1 and -1 infinitely often.

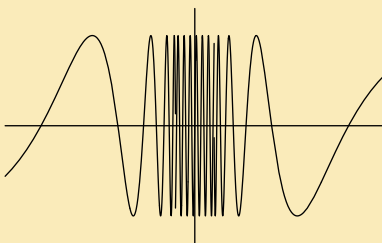


Figure 11.1: Graph of the function $F(x) = \sin x^{-1}$ on $[-\pi/8, \pi/8]$.

Exercise 107, page 26

Consider the function $g(x) = f(x) - x$ which must also be uniformly continuous. Now $g(a) = f(a) - a \geq 0$ and $g(b) = f(b) - b \leq 0$. By the Darboux property there must be a point where $g(c) = 0$. At that point $f(c) - c = 0$.

Exercise 108, page 26

If $z_n \rightarrow c$ then $z_n = f(z_{n-1}) \rightarrow f(c)$. Consequently $c = f(c)$.

Exercise 109, page 26

This is a puzzle. Use the fact that such functions will have maxima and minima in any interval $[c, d] \subset I$ and that continuous functions have the Darboux property.

Exercise 110, page 27

This is again a puzzle. Use the fact that such functions will have maxima and minima in any interval $[c, d] \subset I$ and that continuous functions have the Darboux property. You shouldn't have too much trouble finding an example if I is a closed, bounded interval. What about if I is open?

Exercise 111, page 27

For such functions the one-sided limits $f(x_0+)$ $f(x_0-)$ exist at every point x_0 and $f(x_0-) \leq f(x_0+)$. The function is discontinuous at x_0 if and only if $f(x_0-) < f(x_0+)$. Show that the Darboux property would not allow $f(x_0-) < f(x_0+)$ at any point.

Exercise 112, page 28

In its usual definition

$$F'(x) = \lim_{y \rightarrow x} \frac{F(y) - F(x)}{y - x}$$

or, equivalently,

$$\lim_{y \rightarrow x} \left\{ \frac{F(y) - F(x)}{y - x} - F'(x) \right\} = 0.$$

But limits are defined exactly by ϵ , $\delta(x)$ methods. So that, in fact, this statement about the limit is equivalent to the statement that, for every $\epsilon > 0$ there is a $\delta(x) > 0$ so that

$$\left| \frac{F(y) - F(x)}{y - x} - F'(x) \right| \leq \epsilon$$

whenever $0 < |y - x| < \delta$. [Note the exclusion of $y = x$ here.] This, in turn, is identical to the statement that

$$|F(y) - F(x) - F'(x)(y - x)| \leq \epsilon|y - x|$$

whenever $0 < |y - x| < \delta$. The case $y = x$ which is formally excluded from such statements about limits can be accommodated here because the expression is zero for $y = x$. Consequently, a very small [hardly noticeable] cosmetic change shows that the limit derivative statement is exactly equivalent to the statement that, for every $\epsilon > 0$, there is a $\delta(x) > 0$ so that

$$|F(y) - F(x) - F'(x)(y - x)| \leq \epsilon|y - x|$$

whenever y is a points in I for which $|y - x| < \delta(x)$.

Exercise 113, page 28

As we just proved, if $F'(x_0)$ exists and F is defined on an interval I containing that point then, there is a $\delta(x_0) > 0$ so that

$$|F(y) - F(x_0) - F'(x_0)(y - x)| \leq |y - x_0|$$

whenever y is a point in the interval I for which $|y - x_0| < \delta(x_0)$.

That translates quickly to the statement that

$$|F(y) - F(x_0)| \leq (|F'(x_0)| + 1)|y - x_0|$$

whenever y is a point in the interval I for which $|y - x_0| < \delta(x_0)$.

This gives the clue needed to write up this proof: If $\varepsilon > 0$ then we choose $\delta_1 < \delta(x_0)$ so that

$$\delta_1 < \varepsilon / (|F'(x_0)| + 1).$$

Then

$$|F(y) - F(x_0)| \leq (|F'(x_0)| + 1)|y - x_0| < \varepsilon$$

whenever y is a point in the interval I for which $|y - x_0| < \delta_1$. This is exactly the requirement for continuity at the point x_0 .

Exercise 114, page 29

Note that

$$\begin{aligned} |F(z) - F(y) - F'(x)(z - y)| &= |[F(z) - F(x)] - [F(x) - F(y)] - F'(x)([z - x] - [y - x])| \\ &\leq |F(z) - F(x) - F'(x)(z - x)| + |F(y) - F(x) - F'(x)(y - x)|. \end{aligned}$$

Thus the usual version of this statement quickly leads to the straddled version. Note that the straddled version includes the usual one.

The word “straddled” refers to the fact that, instead of estimating $[F(y) - F(x)]/[y - x]$ to obtain $F'(x)$ we can straddle the point by taking $z \leq x \leq y$ and estimating $[F(y) - F(z)]/[y - z]$, still obtaining $F'(x)$. If we neglect to straddle the point [it would be unstraddled if y and z were on the same side of x] then we would be talking about a much stronger notion of derivative.

Exercise 115, page 29

In preparation . . .

Exercise 116, page 29

If $F'(x_0) > 0$ then, using $\varepsilon = F'(x_0)/2$, there must be a $\delta > 0$ so that

$$|F(z) - F(x_0) - F'(x_0)(z - x_0)| \leq \varepsilon|z - x_0|$$

whenever z is a point in I for which $|z - x_0| < \delta$.

Suppose $x_0 < z < x_0 + \delta$; then it follows from this inequality that

$$F(z) - F(x_0) - F'(x_0)(z - x_0) \geq -\varepsilon(z - x_0)$$

and so

$$F(z) \geq F(x_0) + \varepsilon(z - x_0)/2 > F(x_0).$$

The argument is similar on the left at x_0 .

It is easy to use the definition of locally strictly increasing at a point to show that the derivative is nonnegative. Must it be positive? For a counterexample simply note that $f(x) = x^3$ is locally strictly increasing at every point but that the derivative is not positive everywhere but has a zero at $x_0 = 0$.

Exercise 117, page 29

Take any $[c, d] \subset (a, b)$. We will show that $F(d) > F(c)$ and this will complete the proof that F is strictly increasing on (a, b) .

For each x_0 in $c, d]$ there is a $\delta(x_0) > 0$ so that

$$|F(z) - F(x_0) - F'(x_0)(z - x_0)| \leq \varepsilon|z - x_0|$$

whenever z is a point in I for which $|z - x_0| < \delta(x_0)$.

We apply the Cousin partitioning argument. There must exist at least one partition

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of the interval $[c, d]$ with the property that each interval $[x_{i-1}, x_i]$ has length smaller than $\delta(\xi_i)$. Thus

$$F(d) - F(c) = \sum_{i=1}^n [F(x_i) - F(x_{i-1})] > 0$$

since each of these terms must satisfy

$$[F(x_i) - F(x_{i-1})] = [F(x_i) - F(\xi_i)] + [F(\xi_i) - F(x_{i-1})] > 0.$$

Exercise 118, page 31

The strategy, quite simply, is to argue that there is a point inside the interval where a maximum or minimum occurs. Accordingly the derivative is zero at that point.

First, if f is constant on the interval, then $f'(x) = 0$ for all $x \in (a, b)$, so ξ can be taken to be any point of the interval. Suppose then that f is not constant. Because f is uniformly continuous on the closed, bounded interval $[a, b]$, f achieves a maximum value M and a minimum value m on $[a, b]$.

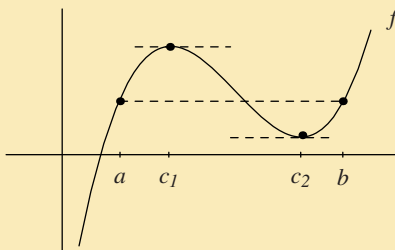
Because f is not constant, one of the values M or m is different from $f(a)$ and $f(b)$, say $M > f(a)$. Choose $f(\xi) = M$. Since $M > f(a) = f(b)$, $c \in (a, b)$. Check that $f'(c) = 0$. If $f'(c) > 0$ then, by Exercise 116, the function f must be locally strictly increasing at x_0 . But this is impossible if x_0 is at a maximum for f . If $f'(c) < 0$ then, again by Exercise 116, the function $-f$ must be locally strictly increasing at x_0 . But this is impossible if x_0 is at a maximum for f . It follows that $f'(c) = 0$.

Exercise 119, page 31

Rolle's theorem asserts that, under our hypotheses, there is a point at which the tangent to the graph of the function is horizontal, and therefore has the same slope as the chord determined by the points $(a, f(a))$ and $(b, f(b))$. (See Figure 11.2.)

Exercise 120, page 31

There may, of course, be many such points; Rolle's theorem just guarantees the existence of at least one such point. You should be able to construct a function, under these hypotheses, with an entire subinterval where the derivative vanishes.

Figure 11.2: Rolle's theorem [note that $f(a) = f(b)$].**Exercise 121, page 31**

First check continuity at every point. This function is not differentiable at zero, but Rolle's theorem requires differentiability only inside the interval, not at the endpoints. Continuity at the point zero is easily checked by using the inequality $|f(x)| = |x \sin x^{-1}| \leq |x|$. Continuity elsewhere follows from the fact that function f is differentiable (by the usual rules) and so continuous. Finally, in order to apply Rolle's theorem, just check that $0 = f(0) = f(1/\pi)$.

There are an infinite number of points between 0 and $1/\pi$ where the derivative is zero. Rolle's theorem guarantees that there is at least one.

Exercise 122, page 31

Yes. Notice that f fails to be differentiable at the endpoints of the interval, but Rolle's theorem does not demand differentiability at either endpoint.

Exercise 123, page 31

No. Notice that f fails to be differentiable only at the midpoint of this interval, but Rolle's theorem demands differentiability at *all* interior points, permitting nondifferentiability only at either endpoint. In this case, even though $f(-1) = f(1)$, there is no point inside the interval where the derivative vanishes.

Exercise 124, page 31

In preparation ...

Exercise 125, page 31

Use Rolle's theorem to show that if x_1 and x_2 are distinct solutions of $p(x) = 0$, then between them is a solution of $p'(x) = 0$.

Exercise 126, page 31

In preparation ...

Exercise 127, page 31

Use Rolle's theorem twice. See Exercise 129 for another variant on the same theme.

Exercise 128, page 31

Since f is continuous we already know (look it up) that f maps $[a, b]$ to some closed, bounded interval $[c, d]$. Use Rolle's theorem to show that there cannot be two values in $[a, b]$ mapping to the same point.

Exercise 129, page 32

cf. Exercise 127.

Exercise 130, page 32

We prove this theorem by subtracting from f a function whose graph is the straight line determined by the chord in question and then applying Rolle's theorem. Let

$$L(x) = f(a) + \frac{f(b) - f(a)}{b - a}(x - a).$$

We see that $L(a) = f(a)$ and $L(b) = f(b)$. Now let

$$g(x) = f(x) - L(x). \quad (11.1)$$

Then g is continuous on $[a, b]$, differentiable on (a, b) , and satisfies the condition $g(a) = g(b) = 0$.

By Rolle's theorem, there exists $c \in (a, b)$ such that $g'(c) = 0$. Differentiating (11.1), we see that $f'(c) = L'(c)$. But

$$L'(c) = \frac{f(b) - f(a)}{b - a},$$

so

$$f'(c) = \frac{f(b) - f(a)}{b - a},$$

as was to be proved.

Exercise 131, page 33

The first statement is just the mean-value theorem applied to every subinterval. For the second statement, note that an increasing function f would allow only positive numbers in S . But increasing functions may have zero derivatives (e.g., $f(x) = x^3$).

Exercise 132, page 33

If t , measured in hours, starts at time $t = 0$ and advances to time $t = 2$ then

$$s'(\tau) = \frac{s(2) - s(0)}{2} = 100/2$$

at some point τ in time between starting and finishing.

Exercise 133, page 33

The mean-value theorem includes Rolle's theorem as a special case. So our previous example $f(x) = \sqrt{|x|}$ which fails to have a derivative at the point $x_0 = 0$ does not satisfy the hypotheses of the mean-value theorem and the conclusion, as we noted earlier, fails.

Exercise 134, page 33

Take any example where the mean-value theorem can be applied and then just change the values of the function at the endpoints.

Exercise 135, page 33

Apply the mean-value theorem to f on the interval $[x, x+a]$ to obtain a point ξ in $[x, x+a]$ with

$$f(x+a) - f(x) = af'(\xi).$$

Exercise 136, page 34

Use the mean-value theorem to compute

$$\lim_{x \rightarrow a+} \frac{f(x) - f(a)}{x - a}.$$

Exercise 137, page 34

This is just a variant on Exercise 136. Show that under these assumptions f' is continuous at x_0 .

Exercise 138, page 34

Use the mean-value theorem to relate

$$\sum_{i=1}^{\infty} (f(i+1) - f(i))$$

to

$$\sum_{i=1}^{\infty} f'(i).$$

Note that f is increasing and treat the former series as a telescoping series.

Exercise 139, page 34

The proof of the mean-value theorem was obtained by applying Rolle's theorem to the function

$$g(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a).$$

For this mean-value theorem apply Rolle's theorem twice to a function of the form

$$h(x) = f(x) - f(a) - f'(a)(x - a) - \alpha(x - a)^2$$

for an appropriate number α .

Exercise 140, page 34

In preparation ...

Exercise 141, page 35

Write

$$\begin{aligned} f(x+h) + f(x-h) - 2f(x) &= \\ [f(x+h) - f(x)] + [f(x-h) - f(x)] \end{aligned}$$

and apply the mean-value theorem to each term.

Exercise 142, page 35

Let

$$\phi(x) = [f(b) - f(a)]g(x) - [g(b) - g(a)]f(x).$$

Then ϕ is continuous on $[a, b]$ and differentiable on (a, b) . Furthermore,

$$\phi(a) = f(b)g(a) - f(a)g(b) = \phi(b).$$

By Rolle's theorem, there exists $\xi \in (a, b)$ for which $\phi'(\xi) = 0$. It is clear that this point ξ satisfies

$$[f(b) - f(a)]g'(\xi) = [g(b) - g(a)]f'(\xi).$$

Exercise 143, page 35

We can interpret the mean-value theorem as applied to curves given parametrically. Suppose f and g are uniformly continuous on $[a, b]$ and differentiable on (a, b) . Consider the curve given parametrically by

$$x = g(t), \quad y = f(t) \quad (t \in [a, b]).$$

As t varies over the interval $[a, b]$, the point (x, y) traces out a curve C joining the points $(g(a), f(a))$ and $(g(b), f(b))$. If $g(a) \neq g(b)$, the slope of the chord determined by these points is

$$\frac{f(b) - f(a)}{g(b) - g(a)}.$$

Cauchy's form of the mean-value theorem asserts that there is a point (x, y) on C at which the tangent is parallel to the chord in question.

Exercise 144, page 35

In its simplest form, l'Hôpital's rule states that for functions f and g , if

$$\lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} g(x) = 0$$

and

$$\lim_{x \rightarrow c} f'(x)/g'(x)$$

exists, then

$$\lim_{x \rightarrow c} \frac{f(x)}{g(x)} = \lim_{x \rightarrow c} \frac{f'(x)}{g'(x)}.$$

You can use Cauchy's mean-value theorem to prove this simple version. Make sure to state your assumptions to match up to the situation in the statement of Cauchy's mean-value theorem.

Exercise 145, page 35

Just expand the determinant.

Exercise 146, page 35

Let $\phi(x)$ be

$$\begin{vmatrix} f(a) & g(a) & h(a) \\ f(b) & g(b) & h(b) \\ f(x) & g(x) & h(x) \end{vmatrix}$$

and imitate the proof of Theorem 142.

Exercise 147, page 36

By the mean-value theorem

$$\frac{f(c) - f(b)}{c - b} = f'(\xi_1) \geq f'(\xi_2) = \frac{f(b) - f(a)}{b - a}$$

for some points $a < \xi_1 < b < \xi_2 < c$. The rest is just elementary algebra.

Note that we should be able to conclude even more if the derivative is strictly increasing since then

$$\frac{f(c) - f(b)}{c - b} = f'(\xi_1) > f'(\xi_2) = \frac{f(b) - f(a)}{b - a}.$$

Exercise 148, page 36

From Lipman Bers, *Classroom Notes: On Avoiding the Mean Value Theorem*, Amer. Math. Monthly 74 (1967), no. 5, 583.

It is hard to agree with this eminent mathematician that students should avoid the mean-value theorem, but (perhaps) for some elementary classes this is reasonable. Here is his proof:

“This is intuitively obvious and easy to prove. Indeed, assume that there is a p , $a < p < b$, such that the set S of all x , $a < x < p$, with $f(x) \geq f(p)$ is not empty. Set $q = \sup S$; since $f'(p) > 0$ we have $a < q < p$. If

$f(q) \geq f(p)$, then since $f'(q) > 0$, there are points of S to the right of q . If $f(q) < f(p)$, then q is not in S and, by continuity, there are no points of S near and to the left of q . Contradiction.

... The “full” mean value theorem, for differentiable but not continuously differentiable functions is a curiosity. It may be discussed together with another curiosity, Darboux’ theorem that every derivative obeys the intermediate value theorem.”

Exercise 149, page 36

From Howard Levi, *Classroom Notes: Integration, Anti-Differentiation and a Converse to the Mean Value Theorem*, Amer. Math. Monthly 74 (1967), no. 5, 585–586.

Exercise 150, page 36

If there are no exceptional points then the usual mean-value theorem does the job. If, say, there is only one point c inside where $f'(c)$ does not exist then apply the mean-value theorem on both of the intervals $[a, c]$ and $[c, b]$ to get two points ξ_1 and ξ_2 so that

$$\frac{f(c) - f(a)}{c - a} = f'(\xi_1)$$

and

$$\frac{f(b) - f(c)}{b - c} = f'(\xi_2).$$

Then

$$|f(b) - f(a)| \leq (c - a)|f'(\xi_1)| + (b - c)|f'(\xi_2)| \leq M(b - a)$$

where for M we just choose whichever is larger, $|f'(\xi_1)|$ or $|f'(\xi_2)|$. A similar proof will handle more exceptional points.

For a method of proof that does not invoke the mean-value theorem see Israel Halperin, *Classroom Notes: A Fundamental Theorem of the Calculus*, Amer. Math. Monthly 61 (1954), no. 2, 122–123.

Exercise 151, page 37

This simple theorem first appears in T. Flett, *A mean value theorem*, Math. Gazette (1958), 42, 38–39.

We can assume that $f'(a) = f'(b) = 0$ [otherwise work with $f(x) - f'(a)x$]. Consider the function $g(x)$ defined to be $[f(x) - f(a)]/[x - a]$ for $x \neq a$ and $f'(a)$ at $x = a$. We compute that

$$g'(x) = -\frac{f(x) - f(a)}{(x-a)^2} + \frac{f'(x)}{x-a} = -\frac{g(x)}{x-a} + \frac{f'(x)}{x-a}.$$

Evidently to prove the theorem is to prove that g' has a zero in (a, b) . Check that such a zero will solve the problem.

To get the zero of g' first consider whether $g(a) = g(b)$. If so then Rolle's theorem does the job. If, instead, $g(b) > g(a)$ then

$$g'(b) = -\frac{g(b)}{b-a} < 0.$$

Thus g is locally decreasing at b . There would then have to be at least one point x_1 for which $g(x_1) > g(b) > g(a)$. The Darboux property of the continuous function g will supply a point x_0 at which $g(x_0) = g(b)$. Apply Rolle's theorem to find that g' has a zero in (x_0, b) . Finally, if $g(b) < g(a)$, then an identical argument should produce the same result.

Exercise 152, page 37

Repeat the arguments for Rolle's theorem with these new hypotheses. Then just take any γ between $F'(a)$ and $F'(b)$ and write $G(x) = F(x) - \gamma x$. If $F'(a) < \gamma < F'(b)$, then $G'(a) = F'(a) - \gamma < 0$ and $G'(b) = F'(b) - \gamma > 0$. This shows then that there is a point $\xi \in (a, b)$ such that $G'(\xi) = 0$. For this ξ we have

$$F'(\xi) = G'(\xi) + \gamma = \gamma,$$

completing the proof for the case $F'(a) < F'(b)$. The proof when $F'(a) > F'(b)$ is similar.

Exercise 153, page 37

If F' is continuous, then it is easy to check that E_α is closed. In the opposite direction suppose that every E_α is closed and F' is not continuous. Then show that there must be a number β and a sequence of points $\{x_n\}$ converging to a point z and yet $f'(x_n) \geq \beta$ and $f'(z) < \beta$. Apply the Darboux property of the derivative to show that this cannot happen if E_β is closed. Deduce that F' is continuous.

Exercise 154, page 37

Polynomials have continuous derivatives and only finitely many points where the value is zero. Let $p(x)$ be a polynomial. Then $p'(x)$ is also a polynomial. Collect all the points $c_1, c_2, \dots, c_p$ where $p'(x) = 0$. In between these points, the value of the derivative is either always positive or always negative otherwise the Darboux property of p' would be violated. On those intervals the function is decreasing or increasing.

Exercise 155, page 38

Take any point $a < x \leq b$ and, applying the mean-value theorem on the interval $[a, x]$, we obtain that

$$|F(x) - F(a)| = F'(\xi)(x - a) = 0(x - a) = 0.$$

Consequently $F(x) = F(a)$ for all $a < x \leq b$. Thus F is constant.

Exercise 156, page 38

In Exercise 148 we established (without the mean-value theorem) that a function with a positive derivative is increasing.

Now we assume that $F'(x) = 0$ everywhere in the interval (a, b) . Consequently, for any integer n , the functions $G(x) = F(x) + x/n$ and $H(x) = x/n - F(x)$ both have a positive derivative and are therefore increasing. In particular, if $x < y$, then

$$H(x) < H(y) \quad \text{and} \quad G(x) < G(y)$$

so that

$$-(y - x)/n < F(y) - F(x) < (y - x)/n$$

would be true for all $n = 1, 2, 3, \dots$. This is only possible if $F(y) = F(x)$.

Exercise 157, page 38

We wish to prove that, if $F : I \rightarrow \mathbb{R}$ is defined at each point of an open interval I and $F'(x) = 0$ for every $x \in I$, then F is a constant function on I . On every closed subinterval $[a, b] \subset I$ the theorem can be applied. Thus f is a constant on the whole interval I . If not then we could find at least two different points x_1 and x_2 with $f(x_1) \neq f(x_2)$. But then we already know that f is constant on the interval $[x_1, x_2]$ (or, rather on the interval $[x_2, x_1]$ if $x_2 < x_1$).

Exercise 158, page 38

Take any points $c < d$ from the interval $[a, b]$ in such a way that (c, d) contains no one of these exceptional points. Consider the closed, bounded interval $[c, d] \subset [a, b]$. An application of the mean-value theorem to this smaller interval shows that

$$F(d) - F(c) = F'(\xi)(d - c) = 0$$

for some point $c < \xi < d$. Thus $F(c) = F(d)$.

Now take any two points $a \leq x_1 < x_2 \leq b$ and find all the exceptional points between them: say $x_1 < c_1 < c_2 < \cdots < c_n < x_2$. On each interval $[x_1, c_1]$, $[c_1, c_2]$, $\dots$, $[c_n, x_2]$ we have (by what we just proved) that

$$F(x_1) = F(c_1), \quad F(c_1) = F(c_2), \quad \dots, \quad F(c_n) = F(x_2).$$

Thus $F(x_1) = F(x_2)$. This is true for any pair of points from the interval $[a, b]$ and so the function is constant.

Exercise 160, page 38

Take any points c and x inside the interval and consider the intervals $[x, c]$ or $[c, x]$. Apply the theorem to determine that F must be constant on any such interval. Consequently $F(x) = F(c)$ for all $a < x < b$.

Exercise 161, page 39

This looks obvious but be a little bit careful with the exceptional set of points x where $F'(x) \neq G'(x)$.

If $F'(x) = f(x)$ for every $a < x < b$ except for points in the finite set C_1 and $G'(x) = f(x)$ for every $a < x < b$ except for points in the finite set C_2 , then the function $H(x) = F(x) - G(x)$ is uniformly continuous on $[a, b]$ and $H'(x) = 0$ for every $a < x < b$ with the possible exception of points in the finite set $C_1 \cup C_2$. By the theorem H is a constant.

Exercise 162, page 39

According to the theorem such a function would have to be discontinuous. Any step function will do here.

Exercise 163, page 39

Let $\varepsilon > 0$. At every point x_0 in the interval (a, b) at which $F'(x_0) = 0$ we can choose a $\delta(x_0) > 0$ so that

$$|F(y) - F(z)| \leq \varepsilon |y - z|$$

for $x_0 - \delta(x_0) < z \leq x_0 \leq y < x_0 + \delta(x_0)$. At the remaining points $a, b, c_1, c_2, c_3, \dots$ we choose $\delta(\cdot)$ so that:

$$\omega(F, [a, a + \delta(a)]) < \frac{\varepsilon}{2},$$

$$\omega(F, [b - \delta(b), b]) < \frac{\varepsilon}{4},$$

and

$$\omega(F, [c_j - \delta(c_j), c_j + \delta(c_j)]) < \frac{\varepsilon}{2^{j+2}}$$

for $j = 1, 2, 3, \dots$. This merely uses the continuity of f at each of these points.

Take any subinterval $[c, d] \subset [a, b]$. By the Cousin covering argument there is a partition

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of the whole interval $[c, d]$ such that

$$\xi_i \in [x_{i-1}, x_i] \text{ and } x_i - x_{i-1} < \delta(\xi_i).$$

For this partition

$$|F(d) - F(c)| \leq \sum_{i=1}^n |F(x_i) - F(x_{i-1})| \leq \varepsilon(d - c) + \sum_{j=1}^{\infty} \frac{\varepsilon}{2^j} = \varepsilon(d - c + 1).$$

This is possible only if $|F(d) - F(c)| = 0$. Since this applies to any such interval $[c, d] \subset [a, b]$ the function must be constant.

Exercise 164, page 39

According to Theorem 1.31 this will be proved if it is possible to write the rational numbers (where $F'(x)$ is not known) as a sequence. This is well-known. To try it on your own. Start off

$$1/1, -1/1, 1/2, -1/2, 2/1, -2/1, 3/1, -3/1, 1/3, -1/3, \dots$$

and describe a listing process that will ultimately include all rational numbers m/n .

Exercise 165, page 39

Apply Exercise 164 to the function $F(x) = G(x) - x^2/2$. Since F is constant, $G(x) = x^2/2 + C$ for some constant C .

Exercise 166, page 39

This looks like an immediate consequence of Theorem 1.31, but we need to be slightly careful about the exceptional sequence of points.

If $F'(x) = f(x)$ for every $a < x < b$ except for points in the sequence $\{c_1, c_2, c_3, \dots\}$ and $G'(x) = f(x)$ for every $a < x < b$ except for points in the sequence $\{d_1, d_2, d_3, \dots\}$, then the function $H(x) = F(x) - G(x)$ is uniformly continuous on $[a, b]$ and $H'(x) = 0$ for every $a < x < b$ with the possible exception of points in the combined sequence $\{c_1, d_1, c_2, d_2, c_3, d_3, \dots\}$. By Theorem 1.31, the function H is a constant.

Exercise 168, page 40

First show directly from the definition that the Lipschitz condition will imply a bounded derivative. Then use the mean-value theorem to get the converse, that is, apply the mean-value theorem to f on the interval $[x, y]$ for any $a \leq x < y \leq b$.

Exercise 169, page 40

The derivative of $f(x) = \sqrt{x} = x^{1/2}$ is the function $f'(x) = x^{1/2}/2$ which exists but is not bounded on $(0, 1)$.

Exercise 170, page 40

One direction is easy. If F is Lipschitz then, for some number M ,

$$|f(x) - f(y)| \leq M|x - y|$$

for all x, y in the interval. In particular

$$\left| \frac{F(x+h) - F(x)}{h} \right| \leq \left| \frac{M|(x+h) - x|}{h} \right| = M.$$

The other direction will take a more sophisticated argument. At each point x_0 choose a $\delta(x_0) > 0$ so that

$$\left| \frac{F(x_0 + h) - F(x_0)}{h} \right| \leq M$$

whenever $x_0 + h \in I$ and $|h| < \delta(x_0)$. Note that, then,

$$\left| \frac{F(y) - F(z)}{y - z} \right| \leq M$$

for $x_0 - \delta(x_0) < z \leq x_0 \leq y < x_0 + \delta(x_0)$. Take any subinterval $[c, d] \subset [a, b]$. By the Cousin partitioning argument there is a partition

$$\{([x_{i-1}, x_i], \xi_i) : i = 1, 2, 3, \dots, n\}$$

of the whole interval $[c, d]$ such that

$$\xi_i \in [x_{i-1}, x_i] \text{ and } x_i - x_{i-1} < \delta(\xi_i).$$

For this partition

$$|F(d) - F(c)| \leq \sum_{i=1}^n |F(x_i) - F(x_{i-1})| \leq \sum_{i=1}^n M|x_i - x_{i-1}| = M(d - c).$$

Thus F is Lipschitz.

Exercise 171, page 40

In preparation ...

Exercise 172, page 40

In preparation ...

Exercise 173, page 40

Yes on any interval (a, ∞) if $a > 0$ but not on $(0, \infty)$.

Exercise 174, page 41

Yes. This is a simple example of a nondifferentiable Lipschitz function, but note that there is only one point of nondifferentiability.

Exercise 177, page 41

From the inequality

$$\left| \frac{F(y) - F(x)}{y - x} \right| \leq M|y - x|$$

deduce that $F'(x) = 0$ everywhere.

Exercise 179, page 41

Find an example illustrating that the first condition can hold without the second condition holding for any value of $K < 1$.

Exercise 180, page 41

Yes if all the functions $F_1, F_2, F_3, \dots$ have the same Lipschitz constant. But, in general, not otherwise.

This is just a simple consequence of the theory of sequence limits and how they behave with inequalities. If we suppose that $x < y$ and that

$$-M(y - x) \leq F_n(y) - F_n(x) \leq M(y - x)$$

for all $n = 1, 2, 3, \dots$, then

$$-M(y - x) \leq \lim_{n \rightarrow \infty} [F_n(y) - F_n(x)] \leq M(y - x)$$

must be true.

Exercise 181, page 44

By the definition, it is indeed an indefinite integral for F' except that we require all indefinite integrals to be continuous. But then we recall that a function is continuous at all points where the derivative exists. So, finally, yes.

Exercise 182, page 44

No. There may be finitely many points where $f(x)$ is not defined, and even if $f(x)$ has been assigned a value there may still be finitely many points where $F'(x) = f(x)$ fails.

Exercise 183, page 44

Exercise 161 is almost identical except that it is stated for two uniformly continuous functions F and G on closed, bounded intervals $[a, b]$. Here (a, b) is open and need not be bounded. But you can apply Exercise 161 to any closed, bounded subinterval of (a, b) .

Exercise 184, page 46

The two functions $F(x) = x$ and $G(x) = 1/x$ are continuous on $(0, 1)$ but G is not uniformly continuous. [It is unbounded and any uniformly continuous function on $(0, 1)$ would have to be bounded.] Thus the two functions $f(x) = F'(x) = 1$ and $g(x) = G'(x) = -1/x^2$ both possess indefinite integrals on the interval $(0, 1)$ so that, of the two indefinite integrals F is uniformly continuous and the other G is not.

Exercise 185, page 46

The mean-value theorem supplies this on any subinterval $[c, d]$ on which F is differentiable; the proof thus requires handling the finite exceptional set. Let M be larger than the values of $|f(x)|$ at points where $F'(x) = f(x)$. Fix $x < y$ in the interval and split the interval at all the points where the derivative F' might not exist:

$$a < x < c_1 < c_2 < \cdots < c_n < y < b.$$

The mean-value theorem supplies that

$$|F(t) - F(s)| \leq M|s - t|$$

on any interval $[s, t]$ for which (s, t) misses all the points of the subdivision. But adding these together we find that

$$|F(t) - F(s)| \leq M|s - t|$$

on any interval $[s, t] \subset [x, y]$. But x and y are completely arbitrary so that

$$|F(t) - F(s)| \leq M|s - t|$$

on any interval $[s, t] \subset (a, b)$.

We already know that if a function is Lipschitz on (a, b) then it is uniformly continuous on (a, b) .

Exercise 186, page 47

It is true that the derivative of $x^3/3 + 1$ is indeed x^2 at every point x . So, provided you also specify the interval in question [here $(-\infty, \infty)$ will do] then the function $F(x) = x^3/3 + 1$ is one possible indefinite integral of $f(x) = x^2$. But there are others and the symbol $\int x^2 dx$ is intended to represent all of them.

Exercise 187, page 47

As we know

$$(x + 1)^2 = x^2 + 2x + 1.$$

The two functions $(x + 1)^2$ and $x^2 + 2x + 1$ differ by a constant (in this case the constant 1). In situations like this it is far better to write

$$\int (2x + 1) dx = (x + 1)^2 + C_1$$

and

$$\int (2x + 1) dx = (x^2 + 2x) + C_2$$

where C_1 and C_2 represent arbitrary constants. Then one won't be using the same letter to represent two different objects.

Exercise 188, page 47

Show that F is an indefinite integral of $f(x) = x^2$ on $(0, 1)$ in this stupid sense if and only if there are three numbers, C_1 , C_2 , and C_3 with $0 < C_1 < 1$ such that $F(x) = x^3/3 + C_2$ on $(0, C_1)$ and $F(x) = x^3/3 + C_3$ on $(C_1, 1)$.

Exercise 189, page 48

By a direct computation

$$\int \frac{1}{x} dx = \log |x| + C$$

on any open interval I for which $0 \notin I$. The function $F(x) = \log |x|$ is a continuous indefinite integral on such an interval I . It cannot be extended to a continuous function on $[0, 1]$ [say] because it is not uniformly continuous. (This is easy to see because uniformly continuous functions are bounded.)

The fact that $f(0) = 1/0$ is undefined is entirely irrelevant. In order for a function to have an indefinite integral [in the calculus sense of this chapter] it is permitted to have finitely many points where it is undefined. See the next exercise where $f(x) = 1/\sqrt{|x|}$ which is also undefined at $x = 0$ but does have an indefinite integral.

Exercise 190, page 48

By a direct computation the function $F(x) = 2\sqrt{x}$ for $x \geq 0$ and $F(x) = -2\sqrt{-x}$ for $x < 0$ has $F'(x) = f(x)$ at every point with the single exception of $x = 0$. Check that this function is continuous everywhere. This is immediate at points where F is differentiable, so it is only at $x = 0$ that one needs to check continuity.

Once again, the fact that $f(0)$ is undefined plays no role in the discussion since this function is defined everywhere else.

Exercise 191, page 48

None of them are correct because no interval is specified. The correct versions would be

$$\int \frac{1}{x} dx = \log x + C \quad \text{on } (0, \infty)$$

or

$$\int \frac{1}{x} dx = \log(-x) + C \quad \text{on } (-\infty, 0)$$

or

$$\int \frac{1}{x} dx = \log |x| + C \quad \text{on } (0, \infty) \text{ or on } (-\infty, 0).$$

You may use subintervals, but we know by now that there are no larger intervals possible.

Exercise 192, page 49

The maximum value of f in each of the intervals $[0, \frac{1}{4}]$, $[\frac{1}{4}, \frac{1}{2}]$, $[\frac{1}{2}, \frac{3}{4}]$, and $[\frac{3}{4}, 1]$ is $1/8$, $1/4$, $9/16$, and 1 respectively. Thus define F to be $x/8$ in the first interval, $1/32 + 1/4(x - 1/4)$ in the second interval, $1/32 + 1/16 + 9/16(x - 1/2)$ in the third interval, and to be $1/32 + 1/16 + 9/64 + (x - 3/4)$ in the final interval. This should be (if the arithmetic was correct) a continuous, piecewise linear function whose slope in each segment exceeds the value of the function f .

Exercise 193, page 49

Start at 0 and first of all work to the right. On the interval $(0, 1)$ the function f has the constant value 1. So define $F(x) = x$ on $[0, 1]$. Then on the interval $(1, 2)$ the function f has the constant value 2. So define $F(x) = 1 + 2(x - 1)$ on $[1, 2]$. Continue until you see how to describe F in general. This is the same construction we used for upper functions.

Exercise 194, page 51

Let F_0 denote the function on $[0, 1]$ that has $F_0(0) = 0$ and has constant slope equal to

$$c_{01} = \sup\{f(t) : 0 < t < 1\}.$$

Subdivide $[0, 1]$ into $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$ and let F_1 denote the continuous, piecewise linear function on $[0, 1]$ that has $F_0(0) = 0$ and has constant slope equal to

$$c_{11} = \sup\{f(t) : 0 < t \leq \frac{1}{2}\}$$

on $[0, \frac{1}{2}]$ and constant slope equal to

$$c_{12} = \sup\{f(t) : \frac{1}{2} \leq t < 1\}$$

on $[0, \frac{1}{2}]$. This construction is continued. For example, at the next stage, Subdivide $[0, 1]$ further into $[0, \frac{1}{4}]$, $[\frac{1}{4}, \frac{1}{2}]$, $[\frac{1}{2}, \frac{3}{4}]$, and $[\frac{3}{4}, 1]$. Let F_2 denote the continuous, piecewise linear function on $[0, 1]$ that has $F_0(0) = 0$ and has constant slope equal to

$$c_{11} = \sup\{f(t) : 0 < t \leq \frac{1}{4}\}$$

on $[0, \frac{1}{4}]$, constant slope equal to

$$c_{12} = \sup\{f(t) : \frac{1}{4} \leq t \leq \frac{1}{2}\}$$

on $[\frac{1}{4}, \frac{1}{2}]$, constant slope equal to

$$c_{13} = \sup\{f(t) : \frac{1}{2} \leq t \leq \frac{3}{4}\}$$

on $[\frac{1}{2}, \frac{3}{4}]$, and constant slope equal to

$$c_{14} = \sup\{f(t) : \frac{3}{4} \leq t < 1\}$$

on $[\frac{3}{4}, 1]$.

In this way we construct a sequence of such functions $\{F_n\}$. Note that each F_n is continuous and nondecreasing. Moreover a look at the geometry reveals that

$$F_n(x) \geq F_{n+1}(x)$$

for all $0 \leq x \leq 1$ and all $n = 0, 1, 2, \dots$. In particular $\{F_n(x)\}$ is a nonincreasing sequence of nonnegative numbers and consequently

$$F(x) = \lim_{n \rightarrow \infty} F_n(x)$$

exists for all $0 \leq x \leq 1$. We prove that $F'(x) = f(x)$ at all points x in $(0, 1)$ at which the function f is continuous.

Fix a point x in $(0, 1)$ at which f is assumed to be continuous and let $\epsilon > 0$. Choose $\delta > 0$ so that the oscillation³

$$\omega f([x - 2\delta, x + 2\delta])$$

of f on the interval $[x - 2\delta, x + 2\delta]$ does not exceed ϵ . Let h be fixed so that $0 < h < \delta$. Choose an integer N sufficiently large that

$$|F_N(x) - F(x)| < \epsilon h \quad \text{and} \quad |F_N(x+h) - F(x+h)| < \epsilon h.$$

From the geometry of our construction notice that the inequality

$$|F_N(x+h) - F_N(x) - f(x)h| \leq h\omega f([x - 2h, x + 2h]),$$

must hold for large enough N . (Simply observe that the graph of F_N will be composed of line segments, each of whose slopes differ from $f(x)$ by no more than the number $\omega f([x - 2h, x + 2h])$.)

Putting these inequalities together we find that

$$\begin{aligned} |F(x+h) - F(x) - f(x)h| &\leq \\ |F_N(x+h) - F_N(x) - f(x)h| + |F_N(x) - F(x)| + |F_N(x+h) - F(x+h)| &< 3\epsilon h. \end{aligned}$$

³See Exercise 61.

This shows that the right-hand derivative of F at x must be exactly $f(x)$. A similar argument will handle the left-hand derivative and we have verified the statement in the theorem about the derivative.

The reader should now check that the function F defined here is Lipschitz on $[0, 1]$. Let M be an upper bound for the function f . Check, first, that

$$0 \leq F_n(y) - F_n(x) \leq M(y - x)$$

for all $x < y$ in $[0, 1]$. Deduce that F is in fact Lipschitz on $[0, 1]$.

Exercise 195, page 51

If $H(t) = G(a + t(b - a))$ then, by the chain rule,

$$H'(t) = G'(a + t(b - a)) \times (b - a) = f(a + t(b - a)) \times (b - a).$$

Substitute $x = a + t(b - a)$ for each $0 \leq t \leq 1$.

Exercise 196, page 51

If $G'(t) = g(t)$ then $\frac{d}{dt}(G(t) + Kt) = g(t) + K = f(t)$.

Exercise 197, page 51

The assumption that f is continuous on an interval (a, b) means that f must be uniformly continuous on any closed subinterval $[c, d] \subset (a, b)$. Such functions are bounded. Applying the theorem gives a continuous function with $F'(x) = f(x)$ everywhere on that interval. This will construct our indefinite integral on (a, b) . Note that F will be Lipschitz on every subinterval $[c, d] \subset (a, b)$ but need not be Lipschitz on (a, b) , because we have not assumed that f is bounded on (a, b) .

Exercise 198, page 52

You need merely to show that H is continuous on the interval and that $H'(x) = rf(x) + sg(x)$ at all but finitely many points in the interval. But both F and G are continuous on that interval and so we need to recall that the sum of continuous functions is again continuous.

Finally we know that $F'(x) = f(x)$ for all x in I except for a finite set C_1 , and we know that $G'(x) = g(x)$ for all x in I except for a finite set C_2 . It follows, by properties of derivatives, that

$$H'(x) = rF'(x) + sG'(x) = rf(x) + sg(x)$$

at all points x in I that are not in the finite set $C_1 \cup C_2$.

Exercise 199, page 52

If F and G both have a derivative at a point x then we know, from the product rule for derivatives, that

$$\frac{d}{dx}[F(x)G(x)] = F'(x)G(x) + F(x)G'(x)$$

Thus, let us suppose that $F'(x)G(x)$ has an indefinite integral $H(x)$ on some interval I . Then H is continuous on I and $H'(x) = F'(x)G(x)$ at all but finitely many points of I . Notice then that

$$K(x) = F(x)G(x) - H(x)$$

satisfies (at points where derivatives exist)

$$K'(x) = F'(x)G(x) + F(x)G'(x) - F'(x)G(x) = F(x)G'(x).$$

Thus

$$\int F(x)G'(x) dx = K(x) + C = F(x)G(x) - H(x) + C = F(x)G(x) - \int F'(x)G(x) dx.$$

Exercise 200, page 53

One memorizes (as a calculus student) the formula

$$\int u dv = uv - \int v du.$$

and makes appropriate substitutions. For example to determine $\int x \cos x dx$ use $u = x$, $v = \cos x dx$, determine $du = dx$ and determine $v = \sin x$ [or $v = \sin x + 1$ for example]. Then substitute in the memorized formula to obtain

$$\int x \cos x dx = \int u dv = uv - \int v du = x \sin x - \int \sin x dx = x \sin x + \cos x + C.$$

or [if you had used $v = \sin x + 1$ instead]

$$\int x \cos x \, dx = x(\sin x + 1) - \int (\sin x + 1) \, dx = x \sin x + \cos x + C.$$

Exercise 201, page 53

If you need more [you are a masochist] then you can find them on the web where we found these. The only reason to spend much further time is if you are shortly to face a calculus exam where some such computation will be required. If you are an advanced student it is enough to remember that “integration by parts” is merely the product rule for derivatives applied to indefinite or definite integration.

There is one thing to keep in mind as a calculus student preparing for questions that exploit integration by parts. An integral can be often split up into many different ways using the substitutions of integration by parts: $u = f(x)$, $dv = g'(x)dx$, $v = g(x)$ and $du = f'(x)dx$. You can do any such problem by trial-and-error and just abandon any unpromising direction. If you care to think in advance about how best to choose the substitution, choose $u = f(x)$ only for functions $f(x)$ that you would care to differentiate and choose $dv = g'(x)dx$ only for functions you would care to integrate.

Exercise 202, page 54

In most calculus courses the rule would be applied only in situations where both functions F and G are everywhere differentiable. For our calculus integral we have been encouraged to permit finitely many exceptional points and to insist then that our indefinite integrals are continuous.

That does not work here: let $F(x) = |x|$ and $G(x) = x^2 \sin x^{-1}$, $G(0) = 0$. Then G is differentiable everywhere and F is continuous with only one point of nondifferentiability. But $F(G(x)) = |x^2 \sin x^{-1}|$ is not differentiable at any point $x = \pm 1/\pi, \pm 1/2\pi, \pm 1/3\pi, \dots$. Thus $F(G(x))$ is not an indefinite integral in the calculus sense for $F'(G(x))$ on any open interval that contains zero and indeed $F'(G(x))$ would have infinitely many points where it is undefined. This function is integrable on any open interval that avoids zero.

This should be considered a limitation of the calculus integral. This is basically an 18th century integral that we are using for teaching purposes. If we allow infinite exceptional sets [as we do in the later integration chapters] then the change of variable rule will hold in great generality.

Exercise 203, page 54

To be precise we should specify an open interval; $(-\infty, \infty)$ will do. To verify the answer itself, just compute

$$\frac{d}{dx} \left\{ \frac{1}{2} \sin(x^2 + 1) \right\} = \cos(x^2 + 1).$$

To verify the steps of the procedure just notice that the substitution $u = x^2 + 1$, $du = 2x dx$ is legitimate on this interval.

Exercise 204, page 54

It would be expected that you have had sufficient experience solving similar problems to realize that integration by parts or other methods will fail but that a change of variable will succeed. The only choices likely in such a simple integral would be $u = x^2$ or perhaps $v = e^{x^2}$. The former leads to

$$\int x e^{x^2} dx = \frac{1}{2} \int e^u du = \frac{1}{2} e^u + C = e^{x^2} + C \quad [u = x^2]$$

since if $u = x^2$ then $du = 2x dx$; the latter leads to

$$\int x e^{x^2} dx = 2 \int v \sqrt{\log v} dv = ? \quad [v = e^{x^2}]$$

since if $v = e^{x^2}$ then $x^2 = \log v$, $x = \sqrt{\log v}$ and $dv = 2x e^{x^2} dx = 2v \sqrt{\log v} dv$.

Note that a wrong choice of substitution may lead to an entirely correct result which does not accord with the instructors "expectation." Usually calculus instructors will select examples that are sufficiently transparent that the correct choice of substitution is immediate. Better yet, they might provide the substitution that they require and ask you to carry it out.

Finally all these computations are valid everywhere so we should state our final result on the interval $(-\infty, \infty)$. Most calculus instructors, however, would not mark you incorrect if you failed to notice this.

Exercise 205, page 54

Assuming that $r \neq 0$ (in which case we are integrating a constant function), use the substitution $u = rx + s$, $du = r dx$ to obtain

$$\int f(rx + s) dx = \frac{1}{r} \int f(u) du = F(u) + C = F(rs + x) + C.$$

This is a linear change of variables and is the most common change of variable in numerous situations.

This can be justified in more detail this way. Suppose that $\int f(t) dx = F(t) + C$ on an open interval I , meaning that F is continuous and $F'(t) = f(t)$ on I with possibly finitely many exceptions. Then find an open interval J so that $rx + s \in I$ for all $x \in J$. It follows that $F(rx + s)$ is continuous on J and that $\frac{d}{dx}F(rx + s) = f(rx + s)$ again with possibly finitely many exceptions. On J then $F(rx + s)$ is an indefinite integral for $f(rx + s)$.

Exercise 206, page 55

This is an exercise in derivatives. Suppose that $f : (a, b) \rightarrow \mathbb{R}$ has an indefinite integral F on the interval (a, b) . Let ξ be a point of continuity of f . We can suppose that ξ is contained in a subinterval $(c, d) \subset (a, b)$ inside which $F'(x) = f(x)$ for all points, except possibly at the point ξ in question.

Let $\varepsilon > 0$. Then, since f is continuous at ξ , there is an interval $[\xi - \delta(\xi), \xi + \delta(\xi)]$ so that

$$f(\xi) - \varepsilon < f(x) < f(\xi) + \varepsilon$$

on that interval. For any $u < \xi < v$ in this smaller interval

$$(f(\xi) - \varepsilon)(\xi - u) \leq F(\xi) - F(u)$$

and

$$(f(\xi) + \varepsilon)(v - \xi) \geq F(v) - F(\xi).$$

This is because the function F is continuous on $[u, \xi]$ and $[\xi, v]$ and has a derivative larger than $(f(\xi) - \varepsilon)$ on (u, ξ) and a derivative smaller than $(f(\xi) + \varepsilon)$ on (ξ, v) . Together these inequalities prove that, for any $u \leq \xi \leq v$, $u \neq v$ on the interval $(\xi - \delta(\xi), \xi + \delta(\xi))$ the inequality

$$\left| \frac{F(v) - F(u)}{v - u} - f(\xi) \right| \leq \varepsilon$$

must be valid. But this says that $F'(\xi) = f(\xi)$.

Exercise 207, page 57

In fact this is particularly sloppy. The function $\log(x-8)$ is defined only for $x > 8$ while $\log(x+5)$ is defined only for $x > -5$. Thus an appropriate interval for the expression given here would be $(8, \infty)$. But it is also true that

$$\int \frac{x+3}{x^2-3x-40} dx = [11/13]\log(|x-8|) + [2/13]\log(|x+5|) + C$$

on any open interval that does not contain either of the points $x = 8$ or $x = -5$. For example, on the interval $(-5, 8)$, the following is valid:

$$\int \frac{x+3}{x^2-3x-40} dx = [11/13]\log(8-x) + [2/13]\log(x+5) + C.$$

You might even prefer to write

$$\int \frac{x+3}{x^2-3x-40} dx = \frac{1}{13} \log \{ (8-x)^{11}(x+5)^2 \} + C.$$

Exercise 208, page 62

Find the necessary statements from Chapter 1 from which this can be concluded.

Exercise 209, page 64

Let us just do the infinite integral. If $\int_{-\infty}^{\infty} f(x) dx$ exists then there is an indefinite integral F on $(-\infty, \infty)$ and both $F(\infty)$ and $F(-\infty)$ exist. By definition

$$\begin{aligned} \int_a^{\infty} f(x) dx &= F(\infty) - F(a), \\ \int_{-\infty}^b f(x) dx &= F(b) - F(-\infty) \end{aligned}$$

and

$$\int_a^b f(x) dx = F(b) - F(a)$$

must all exist.

Exercise 210, page 64

Let us do the additivity formula for the infinite integral. If $\int_{-\infty}^{\infty} f(x) dx$ exists then there is an indefinite integral F on $(-\infty, \infty)$ and both $F(\infty)$ and $F(-\infty)$ exist. By definition

$$\int_{-\infty}^{\infty} f(x) dx = F(\infty) - F(-\infty) = [F(\infty) - F(b)] + [F(b) - F(a)] + [F(a) - F(-\infty)] =$$

$$\int_{-\infty}^a f(x) dx + \int_a^b f(x) dx + \int_b^{\infty} f(x) dx.$$

Any other additivity formula can be proved the same way.

The theorem requires proving another observation. If we know that the integral exists on two abutting intervals then we must check that it exists on the union. Here is the method. If

$$\int_{-\infty}^a f(x) dx \quad \text{and} \quad \int_a^{\infty} f(x) dx$$

both exist then select indefinite integrals F on $(-\infty, a]$ and G on $[a, \infty)$. Define $H(x) = F(x)$ for $x \leq a$ and $H(x) = G(x) - G(a) + F(a)$ for $x > a$. Then H is continuous and must therefore be an indefinite integral for f on $(-\infty, \infty)$. We need to know that the limiting values $H(-\infty)$ and $H(\infty)$ both exist. But $H(-\infty) = F(-\infty)$ and $H(\infty) = G(\infty)$. Thus the integral

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

must exist.

Exercise 211, page 64

We suppose that the two functions f, g are both integrable on a closed, bounded interval $[a, b]$ and that $f(x) \leq g(x)$ for all $x \in [a, b]$. We can If F is an indefinite integral for f on $[a, b]$ and G is an indefinite integral for g on $[a, b]$ then set $H(x) = G(x) - F(x)$ and notice that

$$\frac{d}{dx} H(x) = \frac{d}{dx} [G(x) - F(x)] \geq g(x) - f(x) \geq 0$$

except possibly at the finitely many points where the derivative does not have to agree with the function. But we know that continuous functions with nonnegative derivatives are nondecreasing; the finite number of exceptions does not matter

for this statement. Thus $H(b) - H(a) \geq 0$ and so $[G(b) - G(a)] - [F(b) - F(a)] \geq 0$. Consequently

$$\int_a^b f(x) dx = [F(b) - F(a)] \leq [G(b) - G(a)] = \int_a^b g(x) dx.$$

The details are similar for infinite integrals.

Exercise 212, page 64

We know that

$$\int x^2 dx = x^3/3 + C$$

on any interval. So that, in fact, using [for example] the function $F(x) = x^3/3 + 1$ as an indefinite integral,

$$\int_{-1}^2 x^2 dx = F(2) - F(-1) = [2^3/3 + 1] - [(-1)^3/3 + 1] = 3.$$

Exercise 213, page 64

We know that

$$\int \frac{dx}{x} = \log|x| + C$$

on $(0, \infty)$ and on $(-\infty, 0)$.

In particular we do have a continuous indefinite integral on both of the open intervals $(-1, 0)$ and $(0, 1)$. But this indefinite integral is not uniformly continuous. The easiest clue to this is that the function $\log|x|$ is unbounded on both intervals $(-1, 0)$ and $(0, 1)$.

As for integrability on, say the interval $[-1, 1]$. This is even clearer: there is no antiderivative at all, so the function cannot be integrable by definition.

As to the fact that $f(0)$ is undefined: an integrable function may be undefined at any finite number of points. So this was not an issue and did not need to be discussed.

Finally is this function integrable on the unbounded interval $(\infty, -1]$ or on the unbounded interval $[1, \infty)$? No. Simply

check that neither limit

$$\lim_{x \rightarrow \infty} \log x \text{ or } \lim_{x \rightarrow -\infty} \log(-x)$$

exists.

Exercise 214, page 64

We know that the function $F(x) = 2\sqrt{x}$ is uniformly continuous on $[0, 2]$ and that

$$\frac{d}{dx} 2\sqrt{x} = \frac{1}{\sqrt{x}}$$

for all $x > 0$. Thus this function is integrable on $[0, 2]$ and

$$\int_0^2 \frac{1}{\sqrt{x}} dx = F(2) - F(0) = 2\sqrt{2}.$$

The fact that f is undefined at an endpoint [or any one point for that matter] is no concern to us.

Some calculus course instructors may object here, insisting that the ritual known as “improper integration” needs to be invoked. It does not! We have defined the integral in such a way that this procedure is simply part of the definition. For courses that start with the Riemann integral this procedure would not be allowed since unbounded functions are not Riemann integrable. The function $f(x) = x^{-1/2}$ is unbounded on $(0, 2)$ but this causes us no concern since the definition is only about antiderivatives.

Finally, is this function integrable on $[0, \infty)$? No. The endpoint 0 is no problem but $\lim_{x \rightarrow \infty} 2\sqrt{x}$ does not exist.

Exercise 215, page 64

The simplest method to handle this is to split the problem at 0. If $a < 0 < b$ then

$$\int_a^b 1/\sqrt{|x|} dx = \int_a^0 1/\sqrt{|x|} dx + \int_0^b 1/\sqrt{|x|} dx = \int_a^0 1/\sqrt{-x} dx + \int_0^b 1/\sqrt{x} dx$$

if these two integrals exist. For an indefinite integral of f on $(0, \infty)$ use $F(x) = 2\sqrt{x}$ and for an indefinite integral of f on $(-\infty, 0)$ use $F(x) = -2\sqrt{-x}$.

Thus

$$\int_a^0 1/\sqrt{|x|} dx = F(0) - F(a) = 2\sqrt{a}$$

and

$$\int_0^b 1/\sqrt{|x|} dx = F(b) - F(0) = 2\sqrt{b}.$$

Exercise 216, page 64

In defining an integral on $[a, b]$ as

$$\int_a^b f(x) dx = F(b) - F(a)$$

we have allowed $F'(x) = f(x)$ ($a < x < b$) to fail at a finite number of points, say at $c_1 < c_2 < \dots < c_n$ provided we know that F is continuous at each of these points. We could merely take the separate integrals

$$\int_a^{c_1} f(x) dx, \int_{c_1}^{c_2} f(x) dx, \dots, \int_{c_n}^b f(x) dx$$

and add them together whenever we need to. Thus the integral could be defined with no exceptional set and, for applications, ... we'll add up the pieces that you need.

The calculus integral is only a teaching integral. The modern theory requires a much more general integral and that integral can be obtained by allowing an infinite exceptional set. Thus the training that you are getting by handling the finite exceptional set is really preparing you for the infinite exceptional set. Besides we do get a much better integration theory with our definition, a theory that generalizes quite well to the modern theory.

Another thing to keep in mind: when we pass to an infinite exceptional set we may be unable to “split the interval in pieces.” Indeed, we will eventually allow all of the rational numbers as exceptional points where the derivative may not exist.

Exercise 217, page 65

The derivative of F exists at all points in $(0, 1)$ except at these corners $1/n$, $n = 2, 3, 4, 5, \dots$. If $a > 0$ then the interval $[a, 1]$ contains only finitely many corners. But the interval $(0, 1)$ contains infinitely many corners! Thus F' undefined at infinitely many points of $[0, 1]$ and $F(x)$ is not differentiable at these points.

It is clear that F is continuous at all points inside, since it is piecewise linear. At the endpoint 0 we have $F(0)$ and we have to check that $|F(x) - F(0)|$ is small if x is close to zero. This is easy. So F is uniformly continuous on $[0, 1]$ and the identity

$$\int_a^b F'(x) dx = F(b) - F(a)$$

is true for the calculus integral if $a > 0$. It fails for $a = 0$ only because there are too many points where the derivative fails.

What should we do?

1. Accept that F' is not integrable and not worry about such functions?
2. Wait for a slightly more advanced course where an infinite set of exceptional points is allowed?
3. Immediately demand that the calculus integral accommodate a sequence of exceptional points, not merely a finite set?

We have the resources to do the third of these suggestions. We would have to prove this fact though:

If $F, G : [a, b]$ are uniformly continuous functions, if $F'(x) = f(x)$ for all points in (a, b) except points in some sequence $\{c_n\}$ and if $G'(x) = f(x)$ for all points in (a, b) except points in some sequence $\{d_n\}$, then F and G must differ by a constant.

If we prove that then, immediately, the definition of the calculus integral can be extended to handle this troublesome example. This fact is not too hard to prove, but it is nonetheless much harder than the finite case. Remember the latter uses only the mean-value theorem to find a proof. Accepting sequences of exceptional points will make our simple calculus course just a little bit tougher.

So we stay with the finite case for this chapter and then introduce the infinite case in the next. After all, the calculus integral is just a warm-up integral and is not intended to be the final say in integration theory on the real line.

Exercise 218, page 65

(1). There is no function $F'(x) = 1$ for all x irrational and $F'(x) = 0$ if x is rational, on any interval $[c, d]$. To be an indefinite integral in the calculus sense on an interval $[a, b]$ there must subintervals where F is differentiable. Why is there not? Well derivatives have the Darboux property.

(2) There are too many points where f is not defined. Every interval contains infinitely many rationals.

(3) The only possible indefinite integral is $F(x) = x + k$ for some constant. But then $F'(x) = f(x)$ has too many exceptions: at all the points $x_n = 1/n$, $1 = F'(x_n) \neq f(x_n) = c_n$ unless we had insisted that $c_n = 1$ for all but finitely many of the $\{c_n\}$.

Note: The first two are Lebesgue integrable, but not Riemann integrable. The third is Lebesgue integrable and might be Riemann integrable, depending on whether the sequence $\{c_n\}$ is bounded or not. Thus the calculus integral is quite distinct from these other theories.

Exercise 219, page 65

The function $F(x) = x^{p+1}/(p+1)$ is an antiderivative on $(0, \infty)$ if $p \neq -1$. If $p = -1$ then $F(x) = \log x$ is an antiderivative on $(0, \infty)$. Thus for your answer you will need to check $F(0+)$, $F(1)$, and $F(\infty)$ in all possible cases.

Exercise 220, page 65

Yes. Take an indefinite integral F for f and write

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) dx &= F(\infty) - F(-\infty) = F(\infty) - F(b) + F(b) - F(a) + F(a) - F(-\infty) \\ &= \int_{-\infty}^a f(x) dx + \int_a^b f(x) dx + \int_b^{\infty} f(x) dx. \end{aligned}$$

For the second formula write

$$\begin{aligned} \int_0^{\infty} f(x) dx &= F(\infty) - F(0) = \\ &= F(\infty) - F(N) + [F(N) - F(N-1)] + \dots + [F(2) - F(1)] + [F(1) - F(0)] \end{aligned}$$

$$= F(\infty) - F(N) + \sum_{n=1}^N \int_{n-1}^n f(x) dx.$$

Then, since $\lim_{N \rightarrow \infty} [F(\infty) - F(N)] = 0$, it follows that

$$\int_0^\infty f(x) dx = \lim_{N \rightarrow \infty} \sum_{n=1}^N \int_{n-1}^n f(x) dx = \sum_{n=1}^\infty \int_{n-1}^n f(x) dx.$$

The third formula is similar.

Exercise 221, page 66

Step functions are bounded in every interval $[a, b]$ and have only a finite number of steps, so only a finite number of discontinuities.

Exercise 222, page 66

Differentiable functions are continuous at every point and consequently uniformly continuous on any closed, bounded interval.

Exercise 224, page 66

If $f : (a, b) \rightarrow \mathbb{R}$ is bounded we already know that f is integrable and that the statements here must all be valid. If f is an unbounded function that is continuous at all points of (a, b) then there is a continuous function F on (a, b) for which $F'(x) = f(x)$ for all points there. In particular, F is uniformly continuous on any interval $[c, d] \subset (a, b)$ and so serves as an indefinite integral proving that f is integrable on these subintervals.

In order to claim that f is actually integrable on $[a, b]$ we need to be assured that F can be extended to a uniformly continuous function on all of $[a, b]$. But that is precisely what the conditions

$$\lim_{t \rightarrow a+} \int_t^c f(x) dx \quad \text{and} \quad \lim_{t \rightarrow b-} \int_c^t f(x) dx$$

allow, since they verify that the limits

$$\lim_{t \rightarrow a+} F(t) \quad \text{and} \quad \lim_{t \rightarrow b-} F(t)$$

must both exist. We know that this is both necessary and sufficient in order that F should be extendable to a uniformly continuous function on all of $[a, b]$.

Exercise 225, page 67

We suppose that f has an indefinite integral F on (a, b) . We know that f is integrable on any subinterval $[c, d] \subset (a, b)$ but we cannot claim that f is integrable on all of $[a, b]$ until we check uniform continuity of F .

We assume that g is integrable on $[a, b]$ and construct a proof that f is also integrable on $[a, b]$. Let G be an indefinite integral for g on the open interval (a, b) . We know that G is uniformly continuous because g is integrable.

We check, for $a < s < t < b$ that

$$\left| \int_s^t f(x) dx \right| \leq \int_s^t |g(x)| dx = G(t) - G(s).$$

from which we get that

$$|F(t) - F(s)| \leq G(t) - G(s) \quad \text{for all } a < s < t < c.$$

It follows from an easy ϵ, δ argument that the uniform continuity of F follows from the uniform continuity of G . Consequently f is integrable on $[a, b]$.

For the infinite integral, $\int_a^\infty f(x) dx$ the same argument give us the uniform continuity but does not offer the existence of $F(\infty)$. For that we can use Exercise 63. Since $G(\infty)$ must exist in order for the integral $\int_a^\infty g(x) dx$ to exist, that Exercise 63 shows us, that for all $\epsilon > 0$ there should exist a positive number T so that

$$\omega G((T, \infty)) < \epsilon.$$

But we already know that

$$\omega F((T, \infty)) \leq \omega G((T, \infty)) < \epsilon.$$

A further application of that same exercise shows us that $F(\infty)$ does exist.

Exercise 226, page 67

If f is continuous at all points of (a, b) with the exception of points $a < c_1 < c_2 < \dots < c_m < b$ then we can argue on each interval $[a, c_1]$, $[c_1, c_2]$, $\dots$, $[c_m, b]$. If f is integrable on each of these subintervals of $[a, b]$ then, by the additive property, f must be integrable on $[a, b]$ itself.

We know that f has an indefinite integral on (a, c_1) because f is continuous at each point of that interval. By Theorem 3.9 it follows that f is integrable on $[a, c_1]$. The same argument supplies that f is integrable on each interval $[c_1, c_2], \dots, [c_m, b]$.

Exercise 227, page 67

The method used in the preceding exercise will work. If f is continuous at all points of (a, ∞) with the exception of points $a < c_1 < c_2 < \dots < c_m$ then we can argue on each interval $[a, c_1], [c_1, c_2], \dots, [c_m, \infty)$. If f is integrable on each of these subintervals of $[a, \infty)$ then, by the additive property, f must be integrable on $[a, \infty)$ itself.

We know that f has an indefinite integral on (a, c_1) because f is continuous at each point of that interval. By Theorem 3.9 it follows that f is integrable on $[a, c_1]$. The same argument supplies that f is integrable on each interval $[c_1, c_2], \dots, [c_{m-1}, c_m]$. For the final interval $[c_m, \infty)$ note that f is continuous at every point and so has an indefinite integral on (c_m, ∞) . Now invoke Theorem 3.10 to conclude integrability on $[c_m, \infty)$.

Exercise 228, page 68

Note, first, that all of the integrands are continuous on the interval $(0, \pi/2)$. Using the simple inequality

$$x/2 < \sin x < x \quad (0 < x < \pi/2)$$

we can check that, on the interval $(0, \pi/2)$,

$$\frac{1}{\sqrt{2}} \leq \sqrt{\frac{\sin x}{x}} \leq 1,$$

so that the first integral exists because the integrand is continuous and bounded.

For the next two integrals we observe that

$$\frac{1}{\sqrt{2x}} \leq \sqrt{\frac{\sin x}{x^2}} \leq \frac{1}{\sqrt{x}},$$

and

$$\frac{1}{\sqrt{2x}} \leq \sqrt{\frac{\sin x}{x^3}} \leq \frac{1}{x}.$$

Thus, by the comparison test, one integral exists and the other does not. It is only the integral

$$\int_0^{\pi/2} \sqrt{\frac{\sin x}{x^3}} dx$$

that fails to exist by comparison with the integral

$$\frac{1}{\sqrt{2}} \int_0^{\pi/2} \frac{1}{x} dx.$$

Exercise 229, page 68

The comparison test will handle only the third of these integrals proving that it is integrable. We know that the integrands are continuous on $(0, \infty)$ and so there is an indefinite integral in all cases. The inequality $|\sin x| \leq 1$ shows that

$$\left| \frac{\sin x}{x^2} \right| \leq \frac{1}{x^2}$$

and we know that the integral

$$\int_1^{\infty} \frac{1}{x^2} dx$$

converges. That proves, by the comparison test that

$$\int_1^{\infty} \frac{\sin x}{x^2} dx$$

converges.

To handle the other two cases we would have to compute limits at ∞ to determine convergence. The comparison test does not help.

Exercise 230, page 68

If a nonnegative function $f : (a, b) \rightarrow \mathbb{R}$ has a bounded indefinite integral F on (a, b) , then that function F is evidently nondecreasing. We can claim that f is integrable if and only if we can claim that the limits $F(a+)$ and $F(b-)$ exist. For a nondecreasing function F this is equivalent merely to the observation that F is bounded.

Exercise 231, page 68

In view of the previous exercise we should search for a counterexample that is not nonnegative. Find a bounded function $F : (0, 1)$ that is differentiable but is not uniformly continuous. Try $F(x) = \sin x^{-1}$ and take $f = F'$.

Exercise 232, page 68

The focus of your discussion would have to be on points where the denominator $q(x)$ has a zero. If $[a, b]$ contains no points at which q is zero then the integrand is continuous everywhere (even differentiable) at all points of $[a, b]$ so the function is integrable there.

You will need this distinction. A point x_0 is a *zero* of $q(x)$ if $q(x_0) = 0$. A point x_0 is a zero of $q(x)$ of *order* k ($k = 1, 2, 3, \dots$) if

$$q(x) = (x - x_0)^k h(x)$$

for some polynomial $h(x)$ that does not have a zero at x_0 .

Work on an interval $[x_0, c]$ that contains only the one zero x_0 . For example, if x_0 is a zero of the first order for q , $p(x_0) \neq 0$ and the interval contains no other zeros for p and q then there are positive numbers m and M for which

$$\frac{m}{x - x_0} \leq \frac{p(x)}{q(x)} \leq \frac{M}{x - x_0}$$

on the interval $[x_0, c]$. The comparison test supplies the nonintegrability of the function on this interval.

Do the same at higher order zeros (where you will find the opposite conclusion).

Exercise 233, page 68

Compare with Exercise 232. First consider only the case where $[a, \infty]$ contains no zeros of either $p(x)$ or $q(x)$. Then the integral

$$\int_a^X \frac{p(x)}{q(x)} dx$$

exists and it is only the limiting behavior as $X \rightarrow \infty$ that needs to be investigated. The key idea is that if

$$\frac{m}{x} \leq \left| \frac{p(x)}{q(x)} \right|$$

for some $m > 0$ and all $a \leq x \leq \infty$ then the integral must diverge. Similarly if

$$\frac{M}{x^2} \geq \left| \frac{p(x)}{q(x)} \right|$$

for some $M > 0$ and all $a \leq x \leq \infty$ then the integral must converge.

For a further hint, if you need one, consider the following argument used on the integral

$$\int_1^{\infty} \frac{1+x}{1+x+x^2+x^3} dx.$$

Since

$$\lim_{x \rightarrow \infty} \left| x^2 \times \frac{1+x}{1+x+x^2+x^3} \right| = 1$$

it follows that, for all sufficiently large values of x ,

$$\left| x^2 \times \frac{1+x}{1+x+x^2+x^3} \right| < 2$$

or

$$\left| \frac{1+x}{1+x+x^2+x^3} \right| < \frac{2}{x^2}.$$

Exercise 234, page 68

In Exercise 220 we established the identity

$$\int_1^{\infty} f(x) dx = \sum_{n=2}^{\infty} \int_{n-1}^n f(x) dx,$$

valid if the function f is integrable on $[0, \infty)$. Because f is a nonnegative, decreasing function on $[1, \infty)$ we can see that

$$f(n-1) \geq \int_{n-1}^n f(x) dx \geq f(n).$$

From that we can deduce that the series $\sum_{n=1}^{\infty} f(n)$ converges.

Conversely suppose the series converges. Let

$$F(x) = \int_1^x f(t) dt$$

which is an indefinite integral for f on $(0, \infty)$. The function F is nondecreasing. As before,

$$f(n-1) \geq F(n) - F(n-1) = \int_{n-1}^n f(x) dx \geq f(n).$$

We can deduce from this that if the series $\sum_{n=1}^{\infty} f(n)$ converges then the limit

$$F(\infty) = \lim_{x \rightarrow \infty} F(x)$$

exists. It follows that the integral exists.

Exercise 235, page 68

In Exercise 234 we saw that this would not be possible if the function f is also nondecreasing. That should be the clue as to where to look for a counterexample.

Exercise 236, page 69

In Exercise 234 we saw that this would not be possible if the function f is also nondecreasing. Again that is a clue for finding a counterexample.

Exercise 237, page 69

Uniformly continuous functions are integrable. Here fg is also uniformly continuous.

Exercise 238, page 69

Bounded, continuous functions are integrable. Here fg is also bounded continuous and has discontinuities only at the points where one of f or g is continuous.

Exercise 239, page 69

Take $f(x) = g(x) = \frac{1}{\sqrt{x}}$. Then $f(x)g(x) = \frac{1}{x}$ which we know is not integrable on $[0, 1]$. It is the unboundedness of the functions that causes the difficulty. Clearly some unbounded functions are integrable, but “really big” unbounded functions may not be.

Exercise 240, page 69

If both f and g are continuous functions on $[1, \infty)$ then they must be integrable on every bounded interval. So it is just a delicate matter to arrange for them to be integrable without the product being integrable; this requires attention to the large values.

In every interval $[n, n+1]$ ($n = 1, 2, 3, \dots$) choose f to be a continuous, nonnegative function arranged so that

$$\int_n^{n+1} f(x) dx \leq 1/n^2$$

but

$$\int_n^{n+1} (f(x))^2 dx \geq 1/n.$$

This is just an arithmetic problem in each interval. Then observe that, for $N \leq x \leq N+1$,

$$F(x) = \int_1^x f(t) dt \leq \sum_{i=1}^N \frac{1}{i^2}$$

and

$$G(x) = \int_1^x (f(t))^2 dt \geq \sum_{i=1}^{N-1} \frac{1}{i}.$$

The functions F and G are continuous, nondecreasing functions for which, evidently, $F(\infty)$ exists but $G(\infty)$ does not.

Exercise 241, page 69

By the product rule for derivatives

$$(FG)' = F'G + FG'$$

at all but finitely many points. Thus, since FG is uniformly continuous, the function $FG' + F'G$ is integrable.

Exercise 242, page 70

Yes. Just check the two cases.

Exercise 243, page 70

We know this for $a < b < c$. Make sure to state the assumptions and formulate the thing you want to prove correctly. For example, if $b < a = c$ does it work?

Exercise 244, page 71

For a function $x(t) = t^2$ compute the integral $\int_0^1 x(f) df$. That is perfectly legitimate but will make most mathematicians nauseous. How about using d as a dummy variable: compute $\int_0^1 d^2 dd$? Or use the Greek letter π as a dummy variable: what is $\int_0^1 \sin \pi d \pi$?

Most calculus students are mildly amused by this computation:

$$\int_1^2 \frac{d[\text{cabin}]}{[\text{cabin}]} = \{\log \text{cabin}\}_{\text{cabin}=1}^{\text{cabin}=2} = \log 2.$$

Exercise 245, page 72

That is correct but he is being a jerk. More informative is

$$\int e^{2x} dx = e^{2x}/2 + C,$$

which is valid on any interval. More serious, though, is that the student didn't find an indefinite integral so would be obliged to give some argument about the function $f(x) = e^{2x}$ to convince us that it is indeed integrable. An appeal to continuity would be enough.

Exercise 246, page 72

That is correct but she is not being a jerk. There is no simple formula for any indefinite integral of e^{x^2} other than defining it as an integral as she did here (or perhaps an infinite series). Once again the student would be obliged to give some argument about the function $f(x) = e^{x^2}$ to convince us that it is indeed integrable. An appeal to continuity would be enough.

Exercise 247, page 73

Probably, but the student using the notation should remember that the computation at the ∞ end here is really a limit:

$$\lim_{x \rightarrow \infty} \left[-\frac{2}{\sqrt{x}} \right] = 0.$$

Exercise 248, page 74

Let $F(x) = 0$ for $x \leq 0$ and let $F(x) = x$ for $x > 0$. Then F is continuous everywhere and is differentiable everywhere except at $x = 0$. Consider

$$\int_{-1}^1 F'(x) dx = F(1) - F(-1) = 1$$

and try to find a point ξ where $F'(\xi)(1 - (-1)) = 1$.

Exercise 249, page 74

Let m and M be the minimum and maximum values of the function G . It follows that

$$m \int_a^b \varphi(t) dt \leq \int_a^b G(t) \varphi(t) dt \leq M \int_a^b \varphi(t) dt$$

by monotonicity of the integral. Dividing through by $\int_a^b \varphi(t) dt$ (which we can assume is not zero), we have that

$$m \leq \frac{\int_a^b G(t) \varphi(t) dt}{\int_a^b \varphi(t) dt} \leq M.$$

Since $G(t)$ is continuous, the Darboux property of continuous functions (i.e., the intermediate value theorem) implies that there exists $\xi \in [a, b]$ such that

$$G(x) = \frac{\int_a^b G(t)\varphi(t) dt}{\int_a^b \varphi(t) dt}$$

which completes the proof.

Exercise 252, page 74

In Exercise 229 we avoided looking closely at this important integral but let us do so now.

We need to consider the indefinite integral

$$Si(x) = \int_0^x \frac{\sin t}{t} dt$$

which is known as the *sin integral* function and plays a role in many investigations. Since the functions x^{-1} and $\sin x$ are both continuous on $(0, \infty)$ there is an indefinite integral on $(0, \infty)$. There is no trouble at the left-hand endpoint because the integrand is bounded. Hence the function $Si(x)$ is defined for all $0 \leq x < \infty$.

Our job is simply to show that the limit $Si(\infty)$ exists. It is possible using more advanced methods to evaluate the integral and obtain

$$Si(\infty) = \int_0^\infty \frac{\sin x}{x} dx = \frac{\pi}{2}.$$

To obtain that the limit $Si(\infty)$ exists let us apply the mean-value theorem given as Exercise 251. On any interval $[a, b] \subset (0, \infty)$

$$\int_a^b x^{-1} \sin x dx = \frac{\cos a - \cos \xi}{a} + \frac{\cos \xi - \cos b}{b}$$

for some ξ . Consequently

$$|Si(b) - Si(a)| = \left| \int_a^b \frac{\sin x}{x} dx \right| \leq \frac{2}{a} + \frac{2}{b}.$$

From this we deduce that the oscillation of Si on intervals $[T, \infty)$ is small if T is large, i.e., that

$$\omega Si([T, \infty)) \leq \frac{4}{T} \rightarrow 0$$

as $T \rightarrow \infty$. It follows that $\text{Si}(\infty)$ must exist. This proves that the integral is convergent.

Finally let us show that the function

$$F(x) = \int_0^x \left| \frac{\sin t}{t} \right| dt$$

is unbounded. Then we can conclude that the integral diverges and that the Dirichelet integral is convergent but not absolutely convergent.

To see this take any interval $[2n\pi, (2n+1)\pi]$ on which $\sin x$ is nonnegative. Let us apply the mean-value theorem given as Exercise 250. This will show that

$$\int_{2n\pi}^{(2n+1)\pi} \left| \frac{\sin t}{t} \right| dt \geq \frac{1}{2n\pi}.$$

It follows that for x greater than $(2N+1)\pi$

$$\int_0^x \left| \frac{\sin t}{t} \right| dt \geq \sum_{n=1}^N \int_{2n\pi}^{(2n+1)\pi} \left| \frac{\sin t}{t} \right| dt \geq \sum_{n=1}^N \frac{1}{2n\pi}.$$

Consequently F is unbounded.

Exercise 253, page 77

The choice of midpoint

$$\frac{x_i + x_{i-1}}{2} = \xi_i$$

for the Riemann sum gives a sum

$$= \frac{1}{2} \sum_{i=1}^n (x_i^2 - x_{i-1}^2) = \frac{1}{2} [b^2 - x_{n-1}^2 + x_{n-1}^2 - x_{n-2}^2 + \cdots - a^2] = (b^2 - a^2)/2.$$

To explain why this works you might take the indefinite integral $F(x) = x^2/2$ and check that

$$\frac{F(d) - F(c)}{d - c} = \frac{c + d}{2}$$

so that the mean-value always picks out the midpoint of the interval $[c, d]$ for this very simple function.

Exercise 259, page 78

Just take, first, the points ξ_i^* at which we have the exact identity

$$\int_{x_{i-1}}^{x_i} f(x) dx - f(\xi_i^*)(x_i - x_{i-1}) = 0$$

Then for any other point ξ_i ,

$$\left| \int_{x_{i-1}}^{x_i} f(x) dx - f(\xi_i)(x_i - x_{i-1}) \right| = |f(\xi_i) - f(\xi_i^*)|(x_i - x_{i-1}) \leq \omega f([x_i, x_{i-1}]) (x_i - x_{i-1}).$$

The final comparison with

$$\sum_{i=1}^n \omega f([x_i, x_{i-1}]) (x_i - x_{i-1})$$

follows from this.

To get a good approximation of the integral by Riemann sums it seems that we might need

$$\sum_{i=1}^n \omega f([x_i, x_{i-1}]) (x_i - x_{i-1})$$

to be small. Observe that the pieces in the sum here can be made small if (a) the function is continuous so that the oscillations are small, or (b) points where the function is not continuous occur in intervals $[x_i, x_{i-1}]$ that are small. Loosely then we can make these sums small if the function is mostly continuous, i.e., where it is not continuous can be covered by some small intervals that don't add up to much. The modern statement of this is “the function needs to be continuous almost everywhere.”

Exercise 260, page 80

This is the simplest case to prove since we do not have to fuss at the endpoints or at exceptional points where f is discontinuous.

Let $\epsilon > 0$ and choose $\delta > 0$ so that

$$\omega f([c, d]) < \frac{\epsilon}{(b-a)}$$

whenever $[c, d]$ is a subinterval of $[a, b]$ for which $d - c < \delta$. Note then that if

$$\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\}$$

is a partition of $[a, b]$ with intervals shorter than δ then

$$\sum_{i=1}^n \omega f([x_i, x_{i-1}]) (x_i - x_{i-1}) < \sum_{i=1}^n [\varepsilon / (b - a)] (x_i - x_{i-1}) = \varepsilon.$$

Consequently, by Exercise 259,

$$\begin{aligned} & \left| \int_a^b f(x) dx - \sum_{i=1}^n f(\xi_i) (x_i - x_{i-1}) \right| \\ & \leq \sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f(x) dx - f(\xi_i) (x_i - x_{i-1}) \right| < \varepsilon. \end{aligned}$$

Exercise 261, page 80

You can still use the error estimate in Exercise 259, but will have to handle the endpoints differently than you did in Exercise 260.

Exercise 262, page 80

Add one point c of discontinuity of f in (a, b) and prove that case. [Do for c what you did for the endpoints a and b in Exercise 261.]

Exercise 263, page 80

Once we have selected

$$\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\},$$

a partition of $[a, b]$ with intervals shorter than δ we would be free to move the points ξ_i anywhere within the interval. Thus write the inequality and hold everything fixed except, for one value of i , let $x_{i-1} \leq \xi_i \leq x_i$ vary. That can be used to obtain an upper bound for $|f(\xi)|$ for $x_{i-1} \leq \xi \leq x_i$.

Exercise 264, page 80

Of course we can more easily use the definition of the integral and compute that $\int_0^1 x^2 dx = 1/3 - 0$. This exercise shows that, under certain simple conditions, not merely can we approximate the value of the integral by Riemann sums, we can produce a sequence of numbers which converges to the value of the integral. Simply divide the interval at the points $0, 1/n, 2/n, \dots, (n-1)/n$, and 1 . Take $\xi = i/n$ [the right hand endpoint of the interval]. Then the Riemann sum for this partition is

$$\sum_{i=1}^n \left(\frac{i}{n}\right)^2 \frac{1}{n} = \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 + \dots + n^2}{n^3}.$$

As $n \rightarrow \infty$ this must converge to the value of the integral by Theorem 3.17. The student is advised to find the needed formula for

$$1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 + \dots + N^2.$$

and determine whether the limit is indeed the correct value $1/3$.

Exercise 265, page 80

Determine the value of the integral

$$\int_0^1 x^2 dx$$

in the following way. Let $0 < r < 1$ be fixed. Subdivide the interval $[0, 1]$ by defining the points $x_0 = 0, x_1 = r^{n-1}, x_2 = r^{n-2}, \dots, x_{n-1} = r^{n-(n-1)} = r$, and $x_n = r^{n-(-n)} = 1$. Choose the points $\xi_i \in [x_{i-1}, x_i]$ as the right-hand endpoint of the interval. Then

$$\sum_{i=1}^n \xi_i^2 (x_i - x_{i-1}) = \sum_{i=1}^n (r^{n-i})^2 (r^{n-i} - r^{n-i+1}).$$

Note that for every value of n this is a Riemann sum over subintervals whose length is smaller than $1 - r$.

As $r \rightarrow 1 -$ this must converge to the value of the integral by Theorem 3.17. The student is advised to carry out the evaluation of this limit to determine whether the limit is indeed the correct value $1/3$.

Exercise 267, page 82

The same methods will work for this theorem with a little effort. Obtain, first, an inequality of the form

$$\begin{aligned} & |f(\xi_i)g(\xi_i) - f(\xi_i)g(\xi_i^*)| \\ & \leq M(\omega(f, [x_i, x_{i-1}]) + \omega(g, [x_i, x_{i-1}])). \end{aligned}$$

To obtain this use the simple identity

$$a_1a_2 - b_1b_2 = (a_1 - b_1)a_2 + (a_2 - b_2)b_1$$

and use for M an upper bound of the sum function $|f| + |g|$ which is evidently bounded, since both f and g are bounded.

Exercise 268, page 83

Obtain, first, an inequality of the form

$$\begin{aligned} & \left| f_1(\xi_i)f_2(\xi_i)f_3(\xi_i)\dots f_p(\xi_i) - f_1(\xi_i)f_2(\xi_i^{(2)})f_3(\xi_i^{(3)})\dots f_p(\xi_i^{(p)}) \right| \\ & \leq M[\omega(f_1, [x_i, x_{i-1}]) + \omega(f_2, [x_i, x_{i-1}]) + \omega(f_3, [x_i, x_{i-1}]) + \dots + \omega(f_p, [x_i, x_{i-1}])]. \end{aligned}$$

To obtain this use the simple identity

$$\begin{aligned} & a_1a_2a_3\dots a_p - b_1b_2b_3\dots b_p \\ & = (a_1 - b_1)a_2a_3\dots a_p + (a_2 - b_2)b_1a_3\dots a_p + (a_3 - b_3)b_1b_2a_4\dots a_p + \dots \\ & \quad + (a_p - b_p)b_1b_2b_3\dots b_{p-1} \end{aligned}$$

and use an appropriate M .

Exercise 269, page 83

First note that the function $H(f(x), g(x))$ is defined and bounded. To see this just write

$$|H(f(x), g(x))| \leq M(|f(x)| + |g(x)|)$$

and remember that both f and g are bounded. It is also true that this function is continuous at every point of (a, b) with at most finitely many exceptions. To see this, use the inequality

$$|H(f(x), g(x)) - H(f(x_0), g(x_0))| \leq M(|f(x) - f(x_0)| + |g(x) - g(x_0)|)$$

and the definition of continuity.

Thus the integral $\int_a^b H(f(x), g(x)) dx$ exists as a calculus integral and can be approximated by Riemann sums

$$\sum_{i=1}^n H(f(\xi_i), g(\xi_i))(x_i - x_{i-1}).$$

To complete the proof just make sure that these sums do not differ much from these other similar sums:

$$\sum_{i=1}^n H(f(\xi_i), g(\xi_i^*))(x_i - x_{i-1}).$$

That will follow from the inequality

$$\begin{aligned} & |H(f(\xi_i), g(\xi_i)) - H(f(\xi_i), g(\xi_i^*))| \\ & \leq M|g(\xi_i) - g(\xi_i^*)| \leq M\omega(g, [x_{i-1}, x_i]). \end{aligned}$$

Exercise 270, page 85

Notice, first, that

$$\int_a^b f(x) dx = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f(x) dx.$$

Thus

$$\begin{aligned} \left| \int_a^b f(x) dx - \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \right| &= \left| \sum_{i=1}^n \left\{ \int_{x_{i-1}}^{x_i} f(x) dx - f(\xi_i)(x_i - x_{i-1}) \right\} \right| \\ &\leq \sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f(x) dx - f(\xi_i)(x_i - x_{i-1}) \right| \end{aligned}$$

merely by the triangle inequality.

Exercise 271, page 85

This is the simplest case to prove since we do not have to fuss at the endpoints or at exceptional points where F' may fail to exist. Simply let $\varepsilon > 0$ and choose at each point x a number $\delta(x) > 0$ sufficiently small so that

$$|F(z) - F(y) - f(x)(z - y)| < \frac{\varepsilon(z - y)}{b - a}$$

when $0 < z - y < \delta(x)$ and $y \leq x \leq z$. This is merely the statement $F'(x) = f(x)$ translated into ε, δ language.

Now suppose that we have a partition

$$\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\}$$

of the interval $[a, b]$ with each

$$x_i - x_{i-1} < \delta(\xi_i) \quad \text{and} \quad \xi_i \in [x_{i-1}, x_i].$$

Then, using our estimate on each of the intervals $[x_{i-1}, x_i]$,

$$\begin{aligned} & \left| \int_a^b f(x) dx - \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \right| \\ & \leq \sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f(x) dx - f(\xi_i)(x_i - x_{i-1}) \right| \\ & = \sum_{i=1}^n |[F(x_i) - F(x_{i-1})] - f(\xi_i)(x_i - x_{i-1})| < \frac{\varepsilon}{b - a} \sum_{i=1}^n (x_i - x_{i-1}) = \varepsilon. \end{aligned}$$

Exercise 272, page 85

This is still a simpler case to prove since we do not have to fuss at the endpoints and there is only one exceptional point to worry about, not a finite set of such points.

Let $\varepsilon > 0$ and, at each point $x \neq c$, choose a number $\delta(x) > 0$ sufficiently small so that

$$|F(z) - F(y) - f(x)(z - y)| < \frac{\varepsilon(z - y)}{2(b - a)}$$

when $0 < z - y < \delta(x)$ and $y \leq x \leq z$. This is merely the statement $F'(x) = f(x)$ translated into ϵ, δ language.

At $x = c$ select a positive number $\delta(c) > 0$ so that

$$|F(z) - F(y)| + |f(c)|(z - y) < \epsilon/2$$

when $0 < z - y < \delta(x)$ and $y \leq x \leq z$. This is possible because F is continuous at c so that $|F(z) - F(y)|$ is small if z and y are sufficiently close to c ; the second part is small since $|f(c)|$ is simply a nonnegative number.

Now suppose that we have a partition

$$\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\}$$

of the interval $[a, b]$ with each

$$x_i - x_{i-1} < \delta(\xi_i) \quad \text{and} \quad \xi_i \in [x_{i-1}, x_i].$$

Then, using our estimate on each of the intervals $[x_{i-1}, x_i]$,

$$\begin{aligned} & \sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f(x) dx - f(\xi_i)(x_i - x_{i-1}) \right| \\ &= \sum_{i=1}^n |[F(x_i) - F(x_{i-1})] - f(\xi_i)(x_i - x_{i-1})| < \epsilon/2 + \frac{\epsilon}{b-a} \sum_{i=1}^n (x_i - x_{i-1}) = \epsilon. \end{aligned}$$

Note that we have had to add the $\epsilon/2$ in case it happens that one of the $\xi_i = c$. Otherwise we do not need it.

Exercise 273, page 85

Exercise 260 and Exercise 261 illustrate the method. Just add more points, including the endpoints a and b into the argument.

Let $c_1, c_2, \dots, c_M$ be a finite list containing the endpoints a and b and each of the points in the interval where $F'(x) = f(x)$ fails. Let $\epsilon > 0$ and, at each point $x \neq c_i$, choose a number $\delta(x) > 0$ sufficiently small so that

$$|F(z) - F(y) - f(x)(z - y)| < \frac{\epsilon(z - y)}{2(b - a)}$$

when $0 < z - y < \delta(x)$ and $y \leq x \leq z$. This is merely the statement $F'(x) = f(x)$ translated into ϵ, δ language.

At $x = c_j$ ($j = 1, 2, 3, \dots, M$) select a positive number $\delta(c_j) > 0$ so that

$$\omega(F, [a, b] \cap [c_j - \delta(c_j), c_j + \delta(c_j)]) + \delta(c_j)|f(c)| < \epsilon/2M$$

when $0 < z - y < \delta(x)$ and $y \leq x \leq z$. Thus just uses the continuity of F .

Now suppose that we have a partition

$$\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\}$$

of the interval $[a, b]$ with each

$$x_i - x_{i-1} < \delta(\xi_i) \quad \text{and} \quad \xi_i \in [x_{i-1}, x_i].$$

Note, first, that if $\xi_i = c_j$ for some i and j (which might occur at most M times), then

$$\begin{aligned} & |[F(x_i) - F(x_{i-1})] - f(\xi_i)(x_i - x_{i-1})| \\ & \leq \omega(F, [a, b] \cap [c_j - \delta(c_j), c_j + \delta(c_j)]) + \delta(c_j)|f(c)| < \varepsilon/2M. \end{aligned}$$

At any other point $\xi_i \neq c_j$

$$|F(z) - F(y) - f(x)(z - y)| < \frac{\varepsilon(z - y)}{2(b - a)}.$$

Consequently

$$\begin{aligned} & \sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f(x) dx - f(\xi_i)(x_i - x_{i-1}) \right| \\ & = \sum_{i=1}^n |[F(x_i) - F(x_{i-1})] - f(\xi_i)(x_i - x_{i-1})| < \varepsilon/2 + \frac{\varepsilon}{b - a} \sum_{i=1}^n |(x_i - x_{i-1})| = \varepsilon. \end{aligned}$$

Exercise 274, page 85

The proof is an easy exercise in derivatives. Use F and G for the indefinite integrals of f and g . Let N_0 be the set of points x in (a, b) where $f(x) \leq g(x)$ might fail. Suppose that $F'(x) = f(x)$ except on a finite set N_1 . Suppose that $G'(x) = g(x)$ except on a finite set N_2 .

Then $H = G - F$ has $H'(x) = g(x) - f(x) \geq 0$ except on the finite set $N_0 \cup N_1 \cup N_2$. This set is also finite and, since F and G are uniformly continuous on the interval, so too is H . We now know that if H is uniformly continuous on $[a, b]$ and

$$\frac{d}{dx}H(x) \geq 0$$

for all but finitely many points x in (a, b) ; then $H(x)$ must be nondecreasing on $[a, b]$. Finally then $H(a) \leq H(b)$ shows that $F(a) - F(b) \leq G(b) - G(a)$ and hence that

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

Exercise 279, page 88

If the formula

$$\frac{d}{dx}F(G(x)) = F'(G(x))G'(x)$$

holds everywhere then

$$\int_a^b F'(G(x))G'(x) dx = F(G(b)) - F(G(a)).$$

But we also know that

$$\int_{G(a)}^{G(b)} F'(x) dx = F(G(b)) - F(G(a)).$$

Exercise 280, page 88

That does not work here. If $F(x) = |x|$ and $G(x) = x^2 \sin x^{-1}$, $G(0) = 0$, then G is differentiable everywhere and F is continuous with only one point of nondifferentiability. But $F(G(x)) = |x^2 \sin x^{-1}|$ is not differentiable at any point $x = \pm 1/\pi, \pm 1/2\pi, \pm 1/3\pi, \dots$. Thus $F(G(x))$ is not an indefinite integral in the calculus sense for $F'(G(x))$ on $[0, 1]$ and indeed $F'(G(x))$ would have infinitely many points where it is undefined. This function is, however, integrable on any interval that avoids zero since there would then be only finitely many points at which the continuous function $F(G(x))$ is not differentiable.

This is a feature of the calculus integral. Other integration theories can handle this function.

Exercise 283, page 88

If F' is integrable [calculus sense] on $[a, b]$ then F is continuous there and differentiable at all but finitely many points of the interval. Hence the formula

$$\frac{d}{dx}F(G(x)) = F'(G(x))G'(x)$$

holds everywhere with at most finitely many exceptions. Consequently $F'(G(x))G'(x)$ must be integrable and

$$\int_a^b F'(G(x))G'(x) dx = F(G(b)) - F(G(a)) = \int_{G(a)}^{G(b)} F'(x) dx.$$

Exercise 284, page 89

The integrand is continuous at each point of $(0, \pi^2)$ so the inequality

$$\left| \frac{\cos \sqrt{x}}{\sqrt{x}} \right| \leq \frac{1}{\sqrt{x}}$$

and the comparison test can be used to show that the integral exists.

With $F(u) = \sin u$, $F'(u) = \cos u$, $u = \sqrt{x}$, and $2du = dx/\sqrt{x}$, a change of variables shows that

$$\int_0^{\pi^2} \frac{\cos \sqrt{x}}{\sqrt{x}} dx = 2 \int_0^{\pi} \cos u du = 2 \sin \pi - 2 \sin 0.$$

Integrability also follows from the change of variable formula itself. Take $F(u) = \sin u$ and $G(x) = \sqrt{x}$. Then $G'(x) = 1/2\sqrt{x}$. The function $F(G(x))$ is continuous on $[0, \pi^2]$ and is differentiable at every point of the open interval $(0, \pi^2)$ with a derivative

$$\frac{d}{dx}F(G(x)) = \cos(G(x)) \times G'(x) = \frac{\cos \sqrt{x}}{2\sqrt{x}}.$$

It follows that the integral must exist and that

$$\int_0^{\pi^2} \frac{\cos \sqrt{x}}{2\sqrt{x}} dx = F(G(\pi^2)) - F(G(0)).$$

Exercise 285, page 90

The inequalities

$$-|f(x)| \leq f(x) \leq |f(x)|$$

hold at every point x at which f is defined. Since these functions are assumed to be integrable on the interval $[a, b]$,

$$-\int_a^b |f(x)| dx \leq \int_a^b f(x) dx \leq \int_a^b |f(x)| dx$$

which is exactly what the inequality in the exercise asserts.

Exercise 286, page 90

Observe that

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| = \sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f(x) dx \right| \leq \sum_{i=1}^n \int_{x_{i-1}}^{x_i} |f(x)| dx = \int_a^b |f(x)| dx$$

for all choices of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

Exercise 287, page 90

Define the function $F(x) = x \cos(\frac{\pi}{x})$, $F(0) = 0$ and compute

$$F'(x) = \cos(\pi/x) + (\pi/x) \sin(\pi/x), \quad x \neq 0.$$

Thus F is differentiable everywhere except at $x = 0$ and F is continuous at $x = 0$. To see the latter note that $-|x| \leq F(x) \leq |x|$.

Thus F' has a calculus integral on every interval. Note that $F'(x)$ is continuous everywhere except at $x = 0$ and that it is unbounded on $(0, 1)$.

We show that $|F'|$ is not integrable on $[0, 1]$. It is, however, integrable on any subinterval $[c, d]$ for which $0 < c < d \leq 1$ since F' and hence $|F'|$ are continuous at every point in such an interval.

Take any integer k and consider the points $a_k = 2/(2k+1)$, $b_k = 1/k$ and check that $F(a_k) = 0$ while $F(b_k) =$

$(-1)^k/k$. Observe that

$$0 < a_k < b_k < a_{k-1} < b_{k-1} < \cdots < 1$$

and that

$$\int_{a_k}^{b_k} |F'(x)| dx \geq \left| \int_{a_k}^{b_k} F'(x) dx \right| = |F(b_k) - F(a_k)| = \frac{1}{k}.$$

If $|F'|$ were, in fact, integrable on $[0, 1]$ then, summing n of these pieces, we would have

$$\sum_{k=1}^n \frac{1}{k} \leq \sum_{k=1}^n \int_{a_k}^{b_k} |F'(x)| dx \leq \int_0^1 |F'(x)| dx.$$

This is impossible since $\sum_{k=1}^{\infty} \frac{1}{k} = \infty$.

Note: In the language introduced later, you may wish to observe that F is not a function of bounded variation on $[0, 1]$. There is a close connection between this concept and absolute integrability.

Exercise 288, page 91

You can use the same argument but with different arithmetic. This is the traditional example that illustrates that the calculus integral, which integrates all derivatives, is not contained in the Lebesgue integral. Indefinite Lebesgue integrals, since Lebesgue's integral is an absolute integration method, must be of bounded variation on any interval. In contrast, the function

$$F(x) = x^2 \sin\left(\frac{1}{x^2}\right)$$

is everywhere differentiable but fails to have bounded variation on $[0, 1]$.

Exercise 289, page 91

Since f is continuous on (a, b) with at most finitely many exceptions and is bounded it is integrable. But the same is true for $|f|$, since it too has the same properties. Hence both f and $|f|$ are integrable.

Exercise 290, page 92

Subdivide at any one point x inside (a, b) ,

$$a = x_0 < x_1 = x < x_2 = b.$$

Then

$$|F(x) - F(a)| + |F(x) - F(b)| \leq V(F, [a, b]).$$

Consequently

$$|F(x)| \leq |F(a)| + |F(b)| + V(F, [a, b])$$

offers an upper bound for F on $[a, b]$.

Exercise 291, page 92

If $F : [a, b] \rightarrow \mathbb{R}$ is nondecreasing then $T(x) = F(x) - F(a)$. This is because

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| = \sum_{i=1}^n [F(x_i) - F(x_{i-1})] = F(x) - F(a)$$

for all choices of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = x.$$

If $F : [a, b] \rightarrow \mathbb{R}$ is nonincreasing then $T(x) = F(a) - F(x)$. Putting these together yields that $T(x) = |F(x) - F(a)|$ in both cases.

Exercise 292, page 93

Work on the separate subintervals of $[-\pi, \pi]$ on which $\sin x$ is monotonic. For example, it is nondecreasing on $[-\pi/2, \pi/2]$.

Exercise 293, page 93

You should be able to show that

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})|$$

is either 0 (if none of the points chosen was 0) and is 2 (if one of the points chosen was 0). It follows that $V(F, [-1, 1]) = 2$.

Note that this example illustrates that the computation of the sum

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})|$$

doesn't depend merely on making the points close together, but may depend also on which points get chosen. Later on in Exercise 302 we will see that for continuous functions the sum

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})|$$

will be very close to the variation value $V(F, [a, b])$ if we can choose points very close together. For discontinuous functions, as we see here, we had better consider all points and not miss even one.

Exercise 294, page 93

Simplest to state would be $F(x) = 0$ if x is an irrational number and $F(x) = 1$ if x is a rational number. Explain how to choose points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

so that the sum

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| \geq n.$$

Exercise 295, page 93

Suppose that $F : [a, b] \rightarrow \mathbb{R}$ is Lipschitz with a Lipschitz constant K . Then

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| \leq \sum_{i=1}^n K(x_i - x_{i-1}) = K(b - a)$$

for all choices of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

Thus $V(F, [a, b]) \leq K(b - a)$.

The converse is not true and it is easy to invent a counterexample. Every monotonic function is of bounded variation, and monotonic functions need not be Lipschitz, nor even continuous.

Exercise 296, page 93

To estimate $V(F + G, [a, b])$ consider

$$\sum_{i=1}^n |[F(x_i) + G(x_i)] - [F(x_{i-1}) + G(x_{i-1})]|$$

for all choices of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

By the triangle inequality,

$$\begin{aligned} & \sum_{i=1}^n |[F(x_i) + G(x_i)] - [F(x_{i-1}) + G(x_{i-1})]| \\ & \leq \sum_{i=1}^n |F(x_i) - F(x_{i-1})| + \sum_{i=1}^n |G(x_i) - G(x_{i-1})| \leq V(F, [a, b]) + V(G, [a, b]). \end{aligned}$$

Exercise 297, page 93

Use $F = -G$ and then $F + G$ is a constant and so $V(F + G, [a, b]) = 0$. Thus it is easy to supply an example for which

$$V(F + G, [a, b]) < V(F, [a, b]) + V(G, [a, b]).$$

For exact conditions on when equality might be possible see F. S. Cater, *When total variation is additive*, Proceedings of the American Mathematical Society, Volume 84, No. 4, April 1982.

Exercise 298, page 93

This is a substantial theorem and it is worthwhile making sure to master the methods of proof. Mostly it is just a matter of using the definition and working carefully with inequalities.

(2). T is monotonic, nondecreasing on $[a, b]$.

Take $a \leq x < y \leq b$ and consider computing $V(F, [a, x])$. Take any points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = x$$

Observe that the sum

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| + |F(y) - F(x)| \leq V(F, [a, y]).$$

This this would be true for any such choice of points, it follows that

$$V(F, [a, x]) + |F(y) - F(x)| \leq V(F, [a, y]).$$

Thus $T(x) = V(F, [a, x]) \leq V(F, [a, y]) = T(y)$.

(1). for all $a \leq c < d \leq b$,

$$|F(d) - F(c)| \leq V(F, [c, d]) = T(d) - T(c).$$

The first inequality,

$$|F(d) - F(c)| \leq V(F, [c, d])$$

follows immediately from the definition of what $V(F, [c, d])$ means. The second inequality says this:

$$V(F, [a, d]) = V(F, [a, c]) + V(F, [c, d]) \tag{11.2}$$

and it is this that we must prove.

To prove (11.2) we show first that

$$V(F, [a, d]) \geq V(F, [a, c]) + V(F, [c, d]) \quad (11.3)$$

Let $\varepsilon > 0$ and choose points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = c$$

so that

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| > V(F, [a, c]) - \varepsilon.$$

Then choose points

$$c = x_n < x_{n+1} < x_{n+2} < \cdots < x_{m-1} < x_m = d$$

so that

$$\sum_{i=n+1}^m |F(x_i) - F(x_{i-1})| > V(F, [c, d]) - \varepsilon.$$

Observe that

$$\sum_{i=1}^m |F(x_i) - F(x_{i-1})| \leq V(F, [a, d]).$$

Putting this together now you can conclude that

$$V(F, [a, d]) \geq V(F, [a, c]) + V(F, [c, d]) - 2\varepsilon.$$

Since ε is arbitrary the inequality (11.3) follows.

Now we prove that

$$V(F, [a, d]) \leq V(F, [a, c]) + V(F, [c, d]) \quad (11.4)$$

Choose points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = d$$

so that

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| > V(F, [a, d]) - \varepsilon.$$

We can insist that among the points selected is the point c itself [since that does not make the sum any smaller]. So let us claim that $x_k = c$. Then

$$\sum_{i=1}^k |F(x_i) - F(x_{i-1})| \leq V(F, [a, c])$$

and

$$\sum_{i=k+1}^n |F(x_i) - F(x_{i-1})| \leq V(F, [c, d]).$$

Putting this together now you can conclude that

$$V(F, [a, d]) - \varepsilon < V(F, [a, c]) + V(F, [c, d]).$$

Since ε is arbitrary the inequality (11.4) follows. Finally, then, the inequalities (11.3) and (11.4) verify (11.2).

(3). If F is continuous at a point then so too is T .

We argue just on the right at the point a to claim that if F is continuous at a then $T(a+) = T(a) = 0$. The same argument can be repeated at any point and on either side. The value $T(a+)$ exists since T is monotonic, but it might be positive. Let $\varepsilon > 0$ and choose δ_1 so that

$$|T(x) - T(a+)| < \varepsilon$$

if $a < x < a + \delta_1$. Choose δ_2 , using the continuity of F at a , so that

$$|F(x) - F(a)| < \varepsilon$$

if $a < x < a + \delta_2$. Now take any $a < x < \min\{\delta_1, \delta_2\}$. Choose points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = x$$

so that

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| > T(x) - \varepsilon.$$

Observe that

$$|F(x_1) - F(x_0)| = |F(x_1) - F(a)| < \varepsilon$$

and that

$$\sum_{i=2}^n |F(x_i) - F(x_{i-1})| \leq T(x) - T(x_1) \leq T(x) - T(a+) - [T(x_1) - T(a+)] < 2\varepsilon.$$

Putting these together we can conclude that

$$T(x) < 3\varepsilon$$

for all $a < x < \min\{\delta_1, \delta_2\}$. Thus $T(a+) = 0$.

(4). If F is uniformly continuous on $[a, b]$ then so too is T .

This follows from (3).

(5). If F is continuously differentiable at a point then so too is T and, moreover $T'(x_0) = |F'(x_0)|$.

This statement is not true without the continuity assumption so your proof will have to make use of that assumption. We will assume that F is continuously differentiable at a and conclude that the derivative of T on the right at a exists and is equal to $|F'(a)|$. This means that F is differentiable in some interval containing a and that this derivative is continuous at a .

Let $\varepsilon > 0$ and choose δ so that

$$|F'(a) - F'(x)| < \varepsilon$$

if $a < x < a + \delta$. Now choose points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = x$$

so that

$$T(x) \geq \sum_{i=1}^n |F(x_i) - F(x_{i-1})| > T(x) - \varepsilon(x - a).$$

Apply the mean-value theorem on each of the intervals to obtain

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| = \sum_{i=1}^n |F'(\xi_i)|(x_i - x_{i-1}) = |F'(a)|(x - a) \pm \varepsilon(x - a).$$

We can interpret this to yield that

$$|T(x) - T(a) - |F'(a)|(x - a)| \leq 2\varepsilon(x - a)$$

for all $a < x < a + \delta$. This says precisely that the right-hand derivative of T at a is $|F'(a)|$.

(6). If F is uniformly continuous on $[a, b]$ and continuously differentiable at all but finitely many points in (a, b) then F'

is absolutely integrable and

$$F(x) - F(a) = \int_a^x F'(t) dt \quad \text{and} \quad T(x) = \int_a^x |F'(t)| dt.$$

For F' to be absolutely integrable both F' and $|F'|$ must be integrable. Certainly F' is integrable. The reason that $|F'|$ is integrable is that F' is continuous at all but finitely many points in (a, b) and has for an indefinite integral the uniformly continuous function T . This uses (5).

Exercise 299, page 93

The natural way to do this is to write

$$F(x) = \left(V(F, [a, x]) + \frac{F(x)}{2} \right) - \left(V(F, [a, x]) - \frac{F(x)}{2} \right)$$

in which case this expression is called the *Jordan decomposition*. It is then just a matter of checking that the two parts do in fact express f as the difference of two monotonic, nondecreasing functions. Theorem 3.26 contains all the necessary information.

Exercise 300, page 93

The methods in Exercise 287 can be repeated here. First establish continuity. The only troublesome point is at $x = 0$ and, for that, just notice that $-|x| \leq F(x) \leq |x|$ which can be used to show that F is continuous at $x = 0$.

Then to compute the total variation of take any integer k and consider the points $a_k = 2/(2k+1)$, $b_k = 1/k$ and check that $F(a_k) = 0$ while $F(b_k) = (-1)^k/k$. Observe that

$$0 < a_k < b_k < a_{k-1} < b_{k-1} < \cdots < 1.$$

Consequently

$$\sum_{k=1}^n |F(b_k) - F(a_k)| \leq V(F, [0, 1]).$$

But

$$\sum_{k=1}^n \frac{1}{k} = \sum_{k=1}^n |F(b_k) - F(a_k)|$$

and $\sum_{k=1}^{\infty} \frac{1}{k} = \infty$. It follows that $V(F, [0, 1]) = \infty$.

Exercise 301, page 94

See Gerald A. Heuer, *The derivative of the total variation function*, American Mathematical Monthly, Vol. 78, No. 10 (1971), pp. 1110–1112. For the statement about the variation it is enough to work on $[0, 1]$ since the values on $[-1, 0]$ are symmetrical. For the statement about the derivatives, it is enough to work on the right-hand side at 0, since $F(-x) = -F(x)$. Here is the argument for $r = 2$ from Heuer's article. Note that, on each interval $[2/(2n+1)\pi, 2/(2n-1)\pi]$, the function F vanishes at the endpoints and has a single extreme point x_n where

$$1/n\pi < x_n < 2/(2n-1)\pi.$$

Thus the variation on this interval is $2|F(x_n)|$, and

$$(1/n\pi)^2 = |F(1/n\pi)| < |F(x_n)| < x_n^2 < \{2/[(2n-1)\pi]\}^2.$$

By the integral test for series (page 438)

$$\begin{aligned} 1/n &= \int_n^{\infty} dx/x^2 < \sum_{k=n}^{\infty} 1/k^2 < (\pi^2/2)T(2/[(2n-1)\pi]) < \sum_{k=n}^{\infty} [2/(2k-1)]^2 \\ &< \int_{(2n-3)/2}^{\infty} dx/x^2 = 2/(2n-3). \end{aligned}$$

Then, for $2/[(2n-1)\pi] \leq x \leq 2/[(2n-3)\pi]$ (with $n \geq 3$) we have

$$1/n < (\pi^2/2)T(x) < 2/(2n-5),$$

and hence

$$(2n-3)/n\pi < (1/x)T(x) < (4n-2)/[(2n-5)\pi]$$

It follows that the derivative of T on the right at zero is $2/\pi$. By symmetry the same is true on the left so $T'(0) = 2/\pi$.

Exercise 302, page 94

Exercise 302 shows that continuity would be needed for this result, even if there is only one point of discontinuity.

Choose $\varepsilon > 0$ so that $v + \varepsilon < V(F, [a, b])$. Select points

$$a = y_0 < y_1 < y_2 < \cdots < y_{n-1} < y_n = b$$

so that

$$\sum_{j=1}^k |F(y_j) - F(y_{j-1})| > v + \varepsilon. \quad (11.5)$$

Since F is uniformly continuous on $[a, b]$ there is a $\eta > 0$ so that

$$|F(x) - F(x')| < \frac{\varepsilon}{2(k+1)}$$

whenever $|x - x'| < \eta$.

We are now ready to specify our δ : we choose this smaller than η and also smaller than all the lengths $y_j - y_{j-1}$ for $j = 1, 2, 3, \dots, k$. Now suppose that we have made a choice of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

such that each $x_i - x_{i-1} < \delta$. We shall show that

$$v < \sum_{i=1}^n |F(x_i) - F(x_{i-1})| \leq V(F, [a, b]) \quad (11.6)$$

and this will prove the statement in the exercise.

We can split this sum up into two parts: if an interval (x_{i-1}, x_i) contains any one of the points from the collection

$$a = y_0 < y_1 < y_2 < \cdots < y_{n-1} < y_n = b$$

that we started with, then we will call that interval a black interval. Note that, by our choice of δ , a black interval can contain only one of the y_j points. In fact, if $y_j \in (x_{i-1}, x_i)$ we can make use of the fact that

$$|F(x_i) - F(x_{i-1})| \leq |F(x_i) - F(y_j)| + |F(y_j) - F(x_{i-1})| \leq \frac{\varepsilon}{(k+1)}. \quad (11.7)$$

If (x_{i-1}, x_i) contains none of these points we will call it a white interval. The sum in (11.6) is handled by thinking separately about the white intervals and the black intervals.

Let combine all the x_i 's and all the y_j 's:

$$a = z_0 < z_1 < z_2 < \cdots < z_{n-1} < z_n = b.$$

Note that

$$\sum_{p=1}^m |F(z_p) - F(z_{p-1})| > v + \varepsilon. \quad (11.8)$$

This is because the addition of further points always enlarges the sum or leaves it the same.

The inequality (11.6) now follows by comparing it to (11.8). There are extra white intervals perhaps where a new point has been added, but each of these has been enlarged by adding a single point and the total extra contribution is no more than ε because of (11.7).

Exercise 304, page 94

If F is locally of bounded variation at every point $x \in \mathbb{R}$ then the collection

$$\beta = \{([u, v], w) : w \in [u, v], V(F, [u, v]) < \infty\}$$

is a full cover of the real line. Take any interval $[a, b]$ and choose a partition π of the interval $[a, b]$ so that $\pi \subset \beta$. Then

$$V(F, [a, b]) \leq \sum_{([u, v], w) \in \pi} V(F, [u, v]) < \infty.$$

The converse is immediate.

Exercise 305, page 95

We suppose that f is absolutely integrable on $[a, b]$. Thus $|f|$ is integrable here. Observe, then, that

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| = \sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f(x) dx \right| \leq \sum_{i=1}^n \int_{x_{i-1}}^{x_i} |f(x)| dx = \int_a^b |f(x)| dx$$

for all choices of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

It follows that

$$V(F, [a, b]) \leq \int_a^b |f(x)| dx.$$

Consequently F must be a function of bounded variation and we have established an inequality in one direction for the identity

$$V(F, [a, b]) = \int_a^b |f(x)| dx.$$

Let us prove the opposite direction. Since f and $|f|$ are integrable we may apply the Henstock property (Theorem 3.22) to each of them. Write G for an indefinite integral of $|f|$ and recall that F is an indefinite integral of f .

For every $\varepsilon > 0$ and for each point x in $[a, b]$ there is a $\delta(x) > 0$ so that

$$\begin{aligned} \sum_{i=1}^n |F(x_i) - F(x_{i-1}) - f(\xi_i)(x_i - x_{i-1})| &< \varepsilon \\ \sum_{i=1}^n |G(x_i) - G(x_{i-1}) - |f(\xi_i)|(x_i - x_{i-1})| &< \varepsilon \end{aligned}$$

whenever $\{([x_i, x_{i-1}], \xi_i) : i = 1, 2, \dots, n\}$ is a partition of $[a, b]$ with each

$$x_i - x_{i-1} < \delta(\xi_i) \quad \text{and} \quad \xi_i \in [x_{i-1}, x_i].$$

There must exist one such partition and for that partition

$$\begin{aligned} G(b) - G(a) &= \sum_{i=1}^n G(x_i) - G(x_{i-1}) \leq \varepsilon + \sum_{i=1}^n |f(\xi_i)|(x_i - x_{i-1}) \\ &\leq 2\varepsilon \sum_{i=1}^n |F(x_i) - F(x_{i-1})| \leq V(F, [a, b]) + 2\varepsilon. \end{aligned}$$

It follows, since ε can be any positive number, that

$$\int_a^b |f(x)| dx = G(b) - G(a) \leq V(F, [a, b]).$$

This completes the proof.

Exercise 306, page 95

This is a limited theorem but useful to state and fairly easy to prove given what we now know.

We know that F' is integrable on $[a, b]$; indeed, it is integrable by definition even without the assumption about the

continuity of F' . We also know that, if F' is absolutely integrable, then F would have to be of bounded variation on $[a, b]$. So one direction is clear.

To prove the other direction we suppose that F has bounded variation. Let T be the total variation function of F on $[a, b]$. Then, by Theorem 3.26, T is uniformly continuous on $[a, b]$ and T is differentiable at every point at which F is continuously differentiable, with moreover $T'(x) = |F'(x)|$ at such points. Wherever F' is continuous so too is $|F'|$.

Consequently we have this situation: $T : [a, b] \rightarrow \mathbb{R}$ is a uniformly continuous function that is continuously differentiable at every point in a bounded, open interval (a, b) with possibly finitely many exceptions. Thus $T' = |F'|$ is integrable.

Exercise 307, page 97

The limit function is $f(x) = 1/x$ which is continuous on $(0, \infty)$ but certainly not bounded there.

Exercise 308, page 97

Each of the functions is continuous. Notice that for each $x \in (-1, 1)$, $\lim_{n \rightarrow \infty} f_n(x) = 0$ and yet, for $x \geq 1$, $\lim_{n \rightarrow \infty} f_n(x) = 1$. This is easy to see, but it is instructive to check the details since we can use them later to see what is going wrong in this example. At the right-hand side on the interval $[1, \infty)$ it is clear that $\lim_{n \rightarrow \infty} f_n(x) = 1$.

At the other side, on the interval $(-1, 1)$ the limit is zero. For if $-1 < x_0 < 1$ and $\varepsilon > 0$, let $N \geq \log \varepsilon / \log |x_0|$. Then $|x_0|^N \leq \varepsilon$, so for $n \geq N$

$$|f_n(x_0) - 0| = |x_0|^n < |x_0|^N \leq \varepsilon.$$

Thus

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) = \begin{cases} 0 & \text{if } -1 < x < 1 \\ 1 & \text{if } x \geq 1. \end{cases}$$

The pointwise limit f of the sequence of continuous functions $\{f_n\}$ is discontinuous at $x = 1$. (Figure 3.1 shows the graphs of several of the functions in the sequence just on the interval $[0, 1]$.)

Exercise 310, page 97

The sequence of functions $f_n(x)$ converges to zero on $(-1, 1)$ and to $x - 1$ on $[1, \infty)$. Now $f'_n(x) = x^{n-1}$ on $(-1, 1)$, so by the previous exercise (Exercise 308),

$$\lim_{n \rightarrow \infty} f'_n(x) = \begin{cases} 0 & \text{if } -1 < x < 1 \\ 1 & \text{if } x \geq 1, \end{cases}$$

while the derivative of the limit function, fails to exist at the point $x = 1$. Thus

$$\lim_{n \rightarrow \infty} \frac{d}{dx}(f_n(x)) \neq \frac{d}{dx} \left(\lim_{n \rightarrow \infty} f_n(x) \right)$$

at $x = 1$.

Exercise 312, page 99

The concept of uniform convergence would allow this argument. But interchanging two limiting operations cannot be justified with pointwise convergence. Just because this argument looks plausible does not mean that we are under no obligation to use ϵ, δ type of arguments to try to justify it.

Apparently, though, to verify the continuity of f at x_0 we do need to use two limit operations and be assured that the order of passing to the limits is immaterial.

Exercise 313, page 99

If all the functions f_n had the same upper bound this argument would be valid. But each may have a different upper bound so that the first statement should have been

If each f_n is bounded on an interval I then there must be, by definition, a number M_n so that $|f_n(x)| \leq M_n$ for all x in I .

Exercise 314, page 99

In this exercise we illustrate that an interchange of limit operations may not give a correct result.

For each row m , we have $\lim_{n \rightarrow \infty} S_{mn} = 0$. Do the same thing holding n fixed and letting $m \rightarrow \infty$.

Exercise 317, page 100

We have discussed, briefly, the possibility that there is a sequence that contains every rational numbers. This topic appears in greater detail in Chapter 4.

If f did have a calculus integral there would be a function F such that $F' = f$ at all but finitely many points. There would be at least one interval where f is an exact derivative and yet f does not have the Darboux property since it assumes only the values 0 and 1 (and no values in between).

Exercise 319, page 101

The statements that are defined by inequalities (e.g., bounded, convex) or by equalities (e.g., constant, linear) will not lead to an interchange of two limit operations, and you should expect that they are likely true.

Exercise 320, page 101

As the footnote to the exercise explains, this was Luzin's unfortunate attempt as a young student to understand limits. The professor began by saying "What you say is nonsense." He gave him the example of the double sequence $m/(m+n)$ where the limits as $m \rightarrow \infty$ and $n \rightarrow \infty$ cannot be interchanged and continued by insisting that "permuting two passages to the limit *must not be done*." He concluded with "Give it some thought; you won't get it immediately."

As yet another illustration that some properties are not preserved in the limit, compute the length of the curves in Exercise 320 (Fig. 3.3) and compare with the length of the limiting curve [i.e., the straight line $y = x$].

Exercise 321, page 101

The purpose of the exercise is to lead to the notion of uniform convergence as a stronger alternative to pointwise convergence.

Fix ϵ but let the point x_0 vary. Observe that, when x_0 is relatively small in comparison with ϵ , the number $\log x_0$ is large in absolute value compared with $\log \epsilon$, so relatively small values of n suffice for the inequality $|x_0|^n < \epsilon$. On the

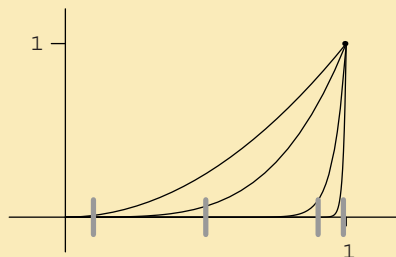


Figure 11.3: The sequence $\{x^n\}$ converges infinitely slowly on $[0, 1]$. The functions $y = x^n$ are shown with $n = 2, 4, 22$, and 100 , with $x_0 = .1, .5, .9$, and $.99$, and with $\varepsilon = .1$.

other hand, when x_0 is near 1 , $\log x_0$ is small in absolute value, so $\log \varepsilon / \log x_0$ will be large. In fact,

$$\lim_{x_0 \rightarrow 1^-} \frac{\log \varepsilon}{\log x_0} = \infty. \quad (11.9)$$

The following table illustrates how large n must be before $|x_0^n| < \varepsilon$ for $\varepsilon = .1$. Note that for $\varepsilon = .1$, there is no single value of N such that $|x_0|^n < \varepsilon$ for every value of $x_0 \in (0, 1)$ and $n > N$. (Figure 11.3 illustrates this.)

x_0	n
.1	2
.5	4
.9	22
.99	230
.999	2,302
.9999	23,025

Some nineteenth-century mathematicians would have described the varying rates of convergence in the example by saying that “the sequence $\{x^n\}$ converges *infinitely slowly* on $(0, 1)$.” Today we would say that this sequence, which does converge pointwise, does *not* converge uniformly. The formulation of the notion of uniform convergence in the next section is designed precisely to avoid this possibility of infinitely slow convergence.

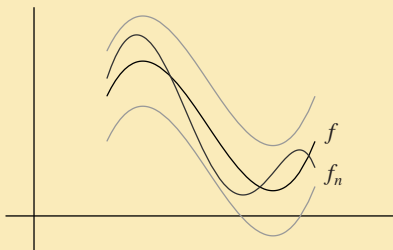


Figure 11.4: Uniform convergence on the whole interval.

Exercise 322, page 102

We observed that the sequence $\{f_n\}$ converges pointwise, but not uniformly, on $(0, 1)$. We realized that the difficulty arises from the fact that the convergence near 1 is very “slow.” But for any fixed η with $0 < \eta < 1$, the convergence is uniform on $[0, \eta]$.

To see this, observe that for $0 \leq x_0 < \eta$, $0 \leq (x_0)^n < \eta^n$. Let $\varepsilon > 0$. Since $\lim_{n \rightarrow \infty} \eta^n = 0$, there exists N such that if $n \geq N$, then $0 < \eta^n < \varepsilon$. Thus, if $n \geq N$, we have

$$0 \leq x_0^n < \eta^n < \varepsilon,$$

so the same N that works for $x = \eta$, also works for all $x \in [0, \eta]$. (See Figure 11.4.)

Exercise 323, page 103

Use the Cauchy criterion for convergence of sequences of real numbers to obtain a candidate for the limit function f . Note that if $\{f_n\}$ is uniformly Cauchy on the interval I , then for each $x \in I$, the sequence of real numbers $\{f_n(x)\}$ is a Cauchy sequence and hence convergent.

Exercise 324, page 103

Fix $n \geq m$ and compute

$$\sup_{x \in [0, \eta]} |x^n - x^m| \leq \eta^m. \quad (11.10)$$

Let $\varepsilon > 0$ and choose an integer N so that $\eta^N < \varepsilon$. Equivalently we require that $N > \log \varepsilon / \log \eta$. Then it follows from (11.10) for all $n \geq m \geq N$ and all $x \in [0, \eta]$ that

$$|x^n - x^m| \leq \eta^m < \varepsilon.$$

We conclude, by the Cauchy criterion, that the sequence $f_n(x) = x^n$ converges uniformly on any interval $[0, \eta]$, for $0 < \eta < 1$. Here there was no computational advantage over the argument in Example 322. Frequently, though, we do not know the limit function and *must* use the Cauchy criterion rather than the definition.

Exercise 325, page 103

This follows immediately from Theorem 3.31. Just check that the translation from series language to sequence language works out in all of the details.

Exercise 326, page 103

Our computations could be based on the fact that the sum of this series is known to us; it is $(1 - x)^{-1}$. We could prove the uniform convergence directly from the definition. Instead let us use the Cauchy criterion.

Fix $n \geq m$ and compute

$$\sup_{x \in [0, \eta]} \left| \sum_{j=m}^n x^j \right| \leq \sup_{x \in [0, \eta]} \left| \frac{x^m}{1 - x} \right| \leq \frac{\eta^m}{1 - \eta}. \quad (11.11)$$

Let $\varepsilon > 0$. Since

$$\eta^m(1 - \eta)^{-1} \rightarrow 0$$

as $m \rightarrow \infty$ we may choose an integer N so that

$$\eta^N(1 - \eta)^{-1} < \varepsilon.$$

Then it follows from (11.11) for all $n \geq m \geq N$ and all $x \in [0, \eta]$ that

$$|x^m + x^{m+1} + \cdots + x^n| \leq \frac{\eta^m}{1 - \eta} < \varepsilon.$$

It follows now, by the Cauchy criterion, that the series converges uniformly on any interval $[0, \eta]$, for $0 < \eta < 1$. Observe, however, that the series does not converge uniformly on $(0, 1)$, though it does converge pointwise there. (See Exercise 341.)

Exercise 327, page 104

It is not always easy to determine whether a sequence of functions is uniformly convergent. In the settings of *series* of functions, this simple test is often useful. This will certainly become one of the most frequently used tools in your study of uniform convergence.

Let $S_n(x) = \sum_{k=0}^n f_k(x)$. We show that $\{S_n\}$ is uniformly Cauchy on I . Let $\varepsilon > 0$. For $m < n$ we have

$$S_n(x) - S_m(x) = f_{m+1}(x) + \cdots + f_n(x),$$

so

$$|S_n(x) - S_m(x)| \leq M_{m+1} + \cdots + M_n.$$

Since the series of constants $\sum_{k=0}^{\infty} M_k$ converges by hypothesis, there exists an integer N such that if $n > m \geq N$,

$$M_{m+1} + \cdots + M_n < \varepsilon.$$

This implies that for $n > m \geq N$,

$$|S_n(x) - S_m(x)| < \varepsilon$$

for all $x \in D$. Thus the sequence $\{S_n\}$ is uniformly convergent on D ; that is, the series $\sum_{k=1}^{\infty} f_k$ is uniformly convergent on I .

Exercise 328, page 104

Then $|x^k| \leq a^k$ for every $k = 0, 1, 2, \dots$ and $x \in [-a, a]$. Since $\sum_{k=0}^{\infty} a^k$ converges, by the M -test the series $\sum_{k=0}^{\infty} x^k$ converges uniformly on $[-a, a]$.

Exercise 329, page 104

The crudest estimate on the size of the terms in this series is obtained just by using the fact that the sine function never exceeds 1 in absolute value. Thus

$$\left| \frac{\sin k\theta}{k^p} \right| \leq \frac{1}{k^p} \quad \text{for all } \theta \in \mathbb{R}.$$

Since the series $\sum_{k=1}^{\infty} 1/k^p$ converges for $p > 1$, we obtain immediately by the M -test that our series converges uniformly (and absolutely) on the interval $(-\infty, \infty)$ [or any interval in fact]. for all real θ provided $p > 1$.

For $0 < p < 1$ the series $\sum_{k=1}^{\infty} 1/k^p$ diverges and the M -test supplies us with no information in these cases.

We seem to have been particularly successful here, but a closer look also reveals a limitation in the method. The series is also pointwise convergent for $0 < p \leq 1$ (use the Dirichlet test) for all values of θ , but it converges nonabsolutely. The M -test cannot be of any help in this situation since it can address only absolutely convergent series. Thus we have obtained only a partial answer because of the limitations of the test.

Because of this observation, it is perhaps best to conclude, when using the M -test, that the series tested “converges absolutely and uniformly” on the set given. This serves, too, to remind us to use a different method for checking uniform convergence of nonabsolutely convergent series. See the next exercise (Exercise 330).

Exercise 330, page 104

We will use the Cauchy criterion applied to the series to obtain uniform convergence. We may assume that the $b_k(x)$ are nonnegative and decrease to zero. Let $\varepsilon > 0$. We need to estimate the sum

$$\left| \sum_{k=m}^n a_k(x) b_k(x) \right| \tag{11.12}$$

for large n and m and all $x \in I$. Since the sequence of functions $\{b_k\}$ converges uniformly to zero on I , we can find an integer N so that for all $k \geq N$ and all $x \in I$

$$0 \leq b_k(x) \leq \frac{\varepsilon}{2M}.$$

The key to estimating the sum (11.12), now, is the summation by parts formula. This is just the elementary identity

$$\sum_{k=m}^n a_k b_k = \sum_{k=m}^n (s_k - s_{k-1}) b_k$$

$$= s_m(b_m - b_{m+1}) + s_{m+1}(b_{m+1} - b_{m+2}) \cdots + s_{n-1}(b_{n-1} - b_n) + s_n b_n.$$

This provides us with

$$\left| \sum_{k=m}^n a_k(x) b_k(x) \right| \leq 2M \left(\sup_{x \in E} |b_m(x)| \right) < \varepsilon$$

for all $n \geq m \geq N$ and all $x \in I$ which is exactly the Cauchy criterion for the series and proves the theorem.

Commentary: The M -test is a highly useful tool for checking the uniform convergence of a series. By its nature, though, it clearly applies only to absolutely convergent series. Abel's test clearly shines in this regard.

It is worth pointing out that in many applications of this theorem the sequence $\{b_k\}$ can be taken as a sequence of numbers, in which case the statement and the conditions that need to be checked are simpler. For reference we can state this as a corollary.

Corollary 11.3 *Let $\{a_k\}$ be a sequence of functions on a set $E \subset \mathbb{R}$. Suppose that there is a number M so that*

$$\left| \sum_{k=1}^N a_k(x) \right| \leq M$$

for all $x \in E$ and every integer N . Suppose that the sequence of real numbers $\{b_k\}$ converges monotonically to zero. Then the series

$$\sum_{k=1}^{\infty} b_k a_k$$

converges uniformly on E .

Exercise 331, page 105

It is possible to prove that this series converges for all θ . Questions about the uniform convergence of this series are intriguing. In Figure 11.5 we have given a graph of some of the partial sums of the series.

The behavior near $\theta = 0$ is most curious. Apparently, if we can avoid that point (more precisely if we can stay a small distance away from that point) we should be able to obtain uniform convergence. Theorem 3.34 will provide a



Figure 11.5: Graph of $\sum_{k=1}^n (\sin k\theta)/k$ on $[0, 2\pi]$ for, clockwise from upper left, $n = 1, 4, 7$, and 10 .

proof. We apply that theorem with $b_k(\theta) = 1/k$ and $a_k(\theta) = \sin k\theta$. All that is required is to obtain an estimate for the sums

$$\left| \sum_{k=1}^n \sin k\theta \right|$$

for all n and all θ in an appropriate set. Let $0 < \eta < \pi/2$ and consider making this estimate on the interval $[\eta, 2\pi - \eta]$. From familiar trigonometric identities we can produce the formula

$$\sin \theta + \sin 2\theta + \sin 3\theta + \sin 4\theta + \cdots + \sin n\theta = \frac{\cos \theta/2 - \cos(2n+1)\theta/2}{2 \sin \theta/2}$$

and using this we can see that

$$\left| \sum_{k=1}^n \sin k\theta \right| \leq \frac{1}{\sin(\eta/2)}.$$

Now Theorem 3.34 immediately shows that

$$\sum_{k=1}^{\infty} \frac{\sin k\theta}{k}$$

converges uniformly on $[\eta, 2\pi - \eta]$.

Figure 11.5 illustrates graphically why the convergence cannot be expected to be uniform near to 0. A computation here is instructive. To check the Cauchy criterion on $[0, \pi]$ we need to show that the sums

$$\sup_{\theta \in [0, \pi]} \left| \sum_{k=m}^n \frac{\sin k\theta}{k} \right|$$

are small for large m, n . But in fact

$$\sup_{\theta \in [0, \pi]} \left| \sum_{k=m}^{2m} \frac{\sin k\theta}{k} \right| \geq \sum_{k=m}^{2m} \frac{\sin(k/2m)}{k} \geq \sum_{k=m}^{2m} \frac{\sin 1/2}{2m} > \frac{\sin 1/2}{2},$$

obtained by checking the value at points $\theta = 1/2m$. Since this is not arbitrarily small, the series cannot converge uniformly on $[0, \pi]$.

Exercise 340, page 106

Use the Cauchy criterion for convergence of sequences of real numbers to obtain a candidate for the limit function f . Note that if $\{f_n\}$ is uniformly Cauchy on a set D , then for each $x \in D$, the sequence of real numbers $\{f_n(x)\}$ is a Cauchy sequence and hence convergent.

Exercise 358, page 111

Let $G_k(x) = \int_0^1 g_k(x) dx$ be the indefinite integrals of the g_k . Observe that, for $k = 2, 3, 4, \dots$, the function G_k is continuous on $[0, 1]$, piecewise linear and that it is differentiable everywhere except at the point $1 - \frac{1}{k}$; it has a right-hand derivative 0 there but a left-hand derivative 2^{-k} . That means that the partial sum

$$F_n(x) = \sum_{k=2}^n g_k(x)$$

is also continuous on $[0, 1]$, piecewise linear and that it is differentiable everywhere except at all the points $1 - \frac{1}{k}$ for $k = 2, 3, 4, \dots$.

Both

$$f(x) = \sum_{n=2}^{\infty} g_n(x) \quad \text{and} \quad F(x) = \sum_{n=2}^{\infty} G_n(x)$$

converge uniformly on $[0, 1]$ and $F'(x) = f(x)$ at every point with the exception of all the points in the sequence $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \dots$. That is too many points for F to be an indefinite integral.

Note that the functions in the sequence $f_1, f_2, f_3, \dots$ are continuous with only finitely many exceptions. But the number of exceptions increase with n . That is the clue that we are heading to a function that may not be integrable in the very severe sense of the calculus integral.

Exercise 359, page 111

Let $\varepsilon > 0$ and choose N so that $|f_n(x) - f(x)| < \varepsilon/(b-a)$ for all $n \geq N$ and all $x \in [a, b]$. Then, since f and each function f_n is integrable,

$$\left| \int_a^b f(x) dx - \int_a^b f_n(x) dx \right| \leq \int_a^b |f(x) - f_n(x)| dx \leq \int_a^b \frac{\varepsilon}{b-a} dx = \varepsilon$$

for all $n \geq N$. This proves that

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

Note that we had to assume that f was integrable in order to make this argument work.

Exercise 360, page 112

Let $g(x) = \lim_n F'_n(x)$. Since each of the functions F'_n is assumed continuous and the convergence is uniform, the function g is also continuous on the interval (a, b) .

From Theorem 3.35 we infer that

$$\int_a^x g(t) dt = \lim_{n \rightarrow \infty} \int_a^x F'_n(t) dt = \lim_{n \rightarrow \infty} [F_n(x) - F_n(a)] = F(x) - F(a) \text{ for all } x \in [a, b]. \quad (11.13)$$

Thus we obtain

$$\int_a^x g(t) dt = F(x) - F(a)$$

or

$$F(x) = \int_a^x g(t) dt + F(a).$$

It follows from the continuity of g that F is differentiable and that $f'(x) = g(x)$ for all $x \in (a, b)$.

Exercise 361, page 112

To justify

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1}$$

we observe first that the series

$$\sum_{k=0}^{\infty} x^k$$

(3.5) converges pointwise on $(-1, 1)$. Next we note (Exercise 342) that the series

$$\sum_{k=1}^{\infty} kx^{k-1}$$

converges pointwise on $(-1, 1)$ and uniformly on any closed interval $[a, b] \subset (-1, 1)$. Thus, if $x \in (-1, 1)$ and $-1 < a < x < b < 1$, then this series converges uniformly on $[a, b]$. Now apply Corollary 3.39.

Indeed there was a bit of trouble on the interval $(-1, 1)$, but trouble that was easily handled by working on a closed, bounded subinterval $[a, b]$ inside.

Exercise 362, page 112

Indeed there is a small bit of trouble on the interval $(-\infty, \infty)$, but trouble that was easily handled by working on a closed, bounded subinterval $[-t, t]$ inside. The Weierstrass M-test can be used to verify uniform convergence since

$$\left| \frac{x^k}{k!} \right| \leq \frac{t^k}{k!}$$

for all $-t < x < t$.

Exercise 366, page 113

The hypotheses of Theorem 3.38 are somewhat more restrictive than necessary for the conclusion to hold and we have relaxed them here by dropping the continuity assumption. That means, though, that we have to work somewhat harder.

We also need not assume that $\{f_n\}$ converges on all of $[a, b]$; convergence at a single point suffices. (We cannot, how-

ever, replace uniform convergence of the sequence $\{f'_n\}$ with pointwise convergence, as Example 310 shows.) Theorem 3.40 applies in a number of cases in which Theorem 3.38 does not.

For the purposes of the proof we can assume that the set of exceptions C is empty. For simply work on subintervals $(c, d) \subset (a, b)$ that miss the set C . After obtaining the proof on each subinterval (c, d) the full statement of the theorem follows by piecing these intervals together.

Let $\varepsilon > 0$. Since the sequence of derivatives converges uniformly on (a, b) , there is an integer N_1 so that

$$|f'_n(x) - f'_m(x)| < \varepsilon$$

for all $n, m \geq N_1$ and all $x \in (a, b)$. Also, since the sequence of numbers $\{f_n(x_0)\}$ converges, there is an integer $N > N_1$ so that

$$|f_n(x_0) - f_m(x_0)| < \varepsilon$$

for all $n, m \geq N$. Let us, for any $x \in [a, b]$, $x \neq x_0$, apply the mean value theorem to the function $f_n - f_m$ on the interval $[x_0, x]$ (or on the interval $[x, x_0]$ if $x < x_0$). This gives us the existence of some point ξ strictly between x and x_0 so that

$$f_n(x) - f_m(x) - [f_n(x_0) - f_m(x_0)] = (x - x_0)[f'_n(\xi) - f'_m(\xi)]. \quad (11.14)$$

From this we deduce that

$$\begin{aligned} |f_n(x) - f_m(x)| &\leq |f_n(x_0) - f_m(x_0)| + |(x - x_0)(f'_n(\xi) - f'_m(\xi))| \\ &< \varepsilon(1 + (b - a)) \end{aligned}$$

for any $n, m \geq N$. Since this N depends only on ε this assertion is true for all $x \in [a, b]$ and we have verified that the sequence of continuous functions $\{f_n\}$ is uniformly Cauchy on $[a, b]$ and hence converges uniformly to a continuous function f on the closed, bounded interval $[a, b]$.

We now know that the one point x_0 where we assumed convergence is any point. Suppose that $a < x_0 < b$. We show that $f'(x_0)$ is the limit of the derivatives $f'_n(x_0)$. Again, for any $\varepsilon > 0$, equation (11.14) implies that

$$|f_n(x) - f_m(x) - [f_n(x_0) - f_m(x_0)]| \leq |x - x_0|\varepsilon \quad (11.15)$$

for all $n, m \geq N$ and any $x \neq x_0$ in the interval (a, b) . In this inequality let $m \rightarrow \infty$ and, remembering that $f_m(x) \rightarrow f(x)$ and $f_m(x_0) \rightarrow f(x_0)$, we obtain

$$|f_n(x) - f_n(x_0) - [f(x) - f(x_0)]| \leq |x - x_0|\varepsilon \quad (11.16)$$

if $n \geq N$. Let C be the limit of the sequence of numbers $\{f'_n(x_0)\}$. Thus there exists $M > N$ such that

$$|f'_M(x_0) - C| < \varepsilon. \quad (11.17)$$

Since the function f_M is differentiable at x_0 , there exists $\delta > 0$ such that if $0 < |x - x_0| < \delta$, then

$$\left| \frac{f_M(x) - f_M(x_0)}{x - x_0} - f'_M(x_0) \right| < \varepsilon. \quad (11.18)$$

From Equation (11.16) and the fact that $M > N$, we have

$$\left| \frac{f_M(x) - f_M(x_0)}{x - x_0} - \frac{f(x) - f(x_0)}{x - x_0} \right| < \varepsilon.$$

This, together with the inequalities (11.17) and (11.18), shows that

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - C \right| < 3\varepsilon$$

for $0 < |x - x_0| < \delta$. This proves that $f'(x_0)$ exists and is the number C , which we recall is $\lim_{n \rightarrow \infty} f'_n(x_0)$.

The final statement of the theorem,

$$\lim_{n \rightarrow \infty} \int_a^b f'_n(x) dx = \int_a^b f'(x) dx,$$

now follows too. We know that f' is the exact derivative on (a, b) of a uniformly continuous function f on $[a, b]$ and so the calculus integral

$$f(b) - f(a) = \int_a^b f'(x) dx.$$

But we also know that

$$f_n(b) - f_n(a) = \int_a^b f'_n(x) dx.$$

and

$$\lim_{n \rightarrow \infty} [f_n(b) - f_n(a)] = f(b) - f(a).$$

Exercise 367, page 113

Let $\{f_k\}$ be a sequence of differentiable functions on an interval $[a, b]$. Suppose that the series $\sum_{k=0}^{\infty} f'_k$ converges uniformly on $[a, b]$. Suppose also that there exists $x_0 \in [a, b]$ such that the series $\sum_{k=0}^{\infty} f_k(x_0)$ converges. Then the series $\sum_{k=0}^{\infty} f_k(x)$ converges uniformly on $[a, b]$ to a function F , F is differentiable, and

$$F'(x) = \sum_{k=0}^{\infty} f'_k(x)$$

for all $a \leq x \leq b$.

Exercise 368, page 113

It is not true. We have already seen a counterexample in Exercise 358.

Here is an analysis of the situation: Let $G_n(x) = \int_a^x g_n(t) dt$. Theorem 3.40 demands a single finite set C of exceptional points where $G'_n(x) = g_n(x)$ might fail. In general, however, this set should depend on n . Thus, for each n select a finite set C_n so that $G'_n(x) = g_n(x)$ is true for all $x \in [a, b] \setminus C_n$.

If $C = \bigcup_{n=1}^{\infty} C_n$ is finite then we could conclude that the limit function g is integrable. But C might be infinite.

Exercise 369, page 114

A simple counterexample, showing that we cannot conclude that $\{f_n\}$ converges on I , is $f_n(x) = n$ for all n . To see there must exist a function f such that $f' = g = \lim_{n \rightarrow \infty} f'_n$ on I : Fix $x_0 \in I$, let $F_n = f_n - f_n(x_0)$ and apply Theorem 3.40 to the sequence $\{F_n\}$. Thus, the uniform limit of a sequence of derivatives $\{f'_n\}$ is a derivative even if the sequence of primitives $\{f_n\}$ does not converge.

Exercise 370, page 116

If there is a finite set of points where one of the inequalities fails redefine all the functions to have value zero there. That cannot change the values of any of the integrals but it makes the inequality valid.

Exercise 371, page 116

Lemma 3.41 is certainly the easier of the two lemmas. For that just notice that, for any integer N , if the inequality

$$f(x) \geq \sum_{k=1}^N g_k(x),$$

holds for all x in (a, b) then, since $\sum_{k=1}^N g_k(x)$ is integrable,

$$\int_a^b f(x) dx \geq \int_a^b \left(\sum_{k=1}^N g_k(x) \right) dx = \sum_{k=1}^N \left(\int_a^b g_k(x) dx \right).$$

But if this inequality in turn is true for all N then

$$\int_a^b f(x) dx \geq \sum_{k=1}^{\infty} \left(\int_a^b g_k(x) dx \right)$$

is also true.

Exercise 372, page 116

This lemma requires a bit of bookkeeping and to make this transparent we will use some language and notation. Because the proof is a bit tricky we will also expand the steps rather more than we usually do.

1. Instead of writing a partition or subpartition out in detail in the form

$$\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$$

we will use the Greek letter⁴ π to denote a partition, so

$$\pi = \{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$$

saves a lot of writing.

2. For the Riemann sum over a partition π , in place of writing the cumbersome

$$\sum_{i=1}^n f(\xi_i)(b_i - a_i)$$

⁴ π is the letter in the Greek alphabet corresponding to “p” so that explains the choice. It shouldn’t interfere with your usual use of this symbol.

we write merely

$$\sum_{([u,v],w) \in \pi} f(w)(v-u) \text{ or } \sum_{\pi} f(w)(v-u)$$

.

3. Instead of saying that a partition satisfies the usual condition

$$\pi = \{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$$

with

$$\xi_i \in [a_i, b_i] \text{ and } b_i - a_i < \delta(\xi_i).$$

we just say π is δ -fine.

This notation will make the arguments transparent and is generally convenient.

Remember that our first step in the proof of Lemma 3.42 is to assume that the inequality

$$f(x) \leq \sum_{k=1}^{\infty} g_k(x),$$

is valid at every point of the interval $[a, b]$. Let $\varepsilon > 0$. Since f itself is assumed to be integrable the interval $[a, b]$, the integral can be approximated (pointwise, not uniformly) by Riemann sums. Thus we can choose, for each $x \in [a, b]$, a $\delta_0(x) > 0$ so that

$$\sum_{\pi} f(w)(v-u) \geq \int_a^b f(x) dx - \varepsilon$$

whenever π is a partition of the interval $[a, b]$ that is δ_0 -fine. This applies Theorem 3.22.

Since g_1 is integrable and, again, the integral can be approximated by Riemann sums we can choose, for each $x \in [a, b]$, a $\delta_0(x) > \delta_1(x) > 0$ so that

$$\sum_{\pi} g_1(w)(v-u) \leq \int_a^b g_1(x) dx + \varepsilon 2^{-1}$$

whenever π is a partition of the interval $[a, b]$ that is δ_1 -fine. Since g_2 is integrable and (yet again) the integral can be approximated by Riemann sums we can choose, for each $x \in [a, b]$, a $\delta_1(x) > \delta_2(x) > 0$ so that

$$\sum_{\pi} g_2(w)(v-u) \leq \int_a^b g_2(x) dx + \varepsilon 2^{-2}$$

whenever π is a partition of the interval $[a, b]$ that is δ_2 -fine. Continuing in this way we find, for each integer $k = 1, 2, 3, \dots$ a $\delta_{k-1}(x) > \delta_k(x) > 0$ so that

$$\sum_{\pi} g_k(w)(v-u) \leq \int_a^b g_k(x) dx + \varepsilon 2^{-k}$$

whenever π is a partition of the interval $[a, b]$ that is δ_k -fine.

Let $t < 1$ and choose for each $x \in [a, b]$ the first integer $N(x)$ so that

$$t f(x) \leq \sum_{n=1}^{N(x)} f_n(x).$$

Let

$$E_n = \{x \in [a, b] : N(x) = n\}.$$

We use these sets to carve up the δ_k and create a new $\delta(x)$. Simply set $\delta(x) = \delta_k(x)$ whenever x belongs to the corresponding set E_k .

Take any partition π of the interval $[a, b]$ that is δ -fine (i.e., it must be a fine partition relative to this newly constructed δ .) The existence of such a partition is guaranteed by the Cousin covering argument. Note that this partition is also δ_0 -fine since $\delta(x) < \delta_0(x)$ for all x . We work carefully with this partition to get our estimates.

Let N be the largest value of $N(w)$ for the finite collection of pairs $([u, v], w) \in \pi$. We need to carve the partition π into a finite number of disjoint subsets by writing, for $j = 1, 2, 3, \dots, N$,

$$\pi_j = \{([u, v], w) \in \pi : w \in E_j\}$$

and

$$\sigma_j = \pi_j \cup \pi_{j+1} \cup \dots \cup \pi_N.$$

for integers $j = 1, 2, 3, \dots, N$. Note that σ_j is itself a subpartition that is δ_j -fine. Putting these together we have

$$\pi = \pi_1 \cup \pi_2 \cup \dots \cup \pi_N.$$

By the way we chose δ_0 and since the new δ is smaller than that we know, for this partition π that

$$\int_a^b f(x) dx - \varepsilon \leq \sum_{\pi} f(w)(v-u)$$

so

$$t \int_a^b f(x) dx - t\varepsilon \leq \sum_{\pi} t f(w)(v - u).$$

We also will remember that for $x \in E_i$,

$$t f(x) \leq g_1(x) + g_2(x) + \cdots + g_i(x).$$

Now we are ready for the crucial computations, each step of which is justified by our observations above:

$$\begin{aligned} t \int_a^b f(x) dx - t\varepsilon &\leq \\ &\sum_{\pi} t f(w)(v - u) = \sum_{i=1}^N \sum_{\pi_i} t f(w)(v - u) \\ &\leq \sum_{i=1}^N \sum_{\pi_i} (g_1(w) + g_2(w) + \cdots + g_i(w))(v - u) \\ &= \sum_{j=1}^N \left(\sum_{\sigma_j} g_j(w)(v - u) \right) \leq \\ &\sum_{j=1}^N \left(\int_a^b g_j(x) dx + \varepsilon 2^{-j} \right) \leq \sum_{j=1}^{\infty} \left(\int_a^b g_j(x) dx \right) + \varepsilon. \end{aligned}$$

Since ε is arbitrary, this shows that

$$t \int_a^b f(x) dx \leq \sum_{k=1}^{\infty} \left(\int_a^b g_k(x) dx \right).$$

As this is true for all $t < 1$ the inequality of the lemma must follow too.

Exercise 374, page 116

This follows from these lemmas and the identity

$$f(x) = f_1(x) + \sum_{n=1}^{\infty} (f_n(x) - f_{n-1}(x)).$$

Since f_n is a nondecreasing sequence of functions, the sequence of functions $f_n(x) - f_{n-1}(x)$ is nonnegative. As usual, ignore the finite set of exceptional points or assume that all functions are set equal to zero at those points.

Exercise 375, page 116

We use the same technique and the same language as used in the solution of Exercise 372.

Let $g_n = f - f_n$ and let G_n denote the indefinite integral of the function g_n . The sequence of functions $\{g_n\}$ is nonnegative and monotone decreasing with $\lim_{n \rightarrow \infty} g_n(x) = 0$ at each x .

Let $\varepsilon > 0$. Choose a sequence of functions $\{\delta_k\}$ so that

$$\sum_{([u,v],w) \in \pi} |G_k(v) - G_k(u) - g_k(w)(v-u)| < \varepsilon 2^{-k}$$

whenever π is a partition of the interval $[a, b]$ that is δ_k -fine. Choose, for each $x \in [a, b]$, the first integer $N(x)$ so that

$$g_k(x) < \varepsilon \text{ for all } k \geq N(x).$$

Let

$$E_n = \{x \in [a, b] : N(x) = n\}.$$

We use these sets to carve up the δ_k and create a new $\delta(x)$. Simply set $\delta(x) = \delta_k(x)$ whenever x belongs to the corresponding set E_k .

Take any partition π of the interval $[a, b]$ that is δ -fine (i.e., it must be a fine partition relative to this newly constructed δ .) The existence of such a partition is guaranteed by the Cousin covering argument.

Let N be the largest value of $N(w)$ for the finite collection of pairs $([u, v], w) \in \pi$. We need to carve the partition π into a finite number of disjoint subsets by writing

$$\pi_j = \{([u, v], w) \in \pi : w \in E_j\}$$

for integers $j = 1, 2, 3, \dots, N$. Note that

$$\pi = \pi_1 \cup \pi_2 \cup \dots \cup \pi_N$$

and that these collections are pairwise disjoint.

Now let m be any integer greater than N . We compute

$$\begin{aligned} 0 \leq \int_a^b g_m(x) dx &= G_m(b) - G_m(a) = \sum_{([u,v],w) \in \pi} (G_m(v) - G_m(u)) = \\ \sum_{j=1}^N \left(\sum_{([u,v],w) \in \pi_j} (G_m(v) - G_m(u)) \right) &\leq \sum_{j=1}^N \left(\sum_{([u,v],w) \in \pi_j} (G_j(v) - G_j(u)) \right) \leq \\ \sum_{j=1}^N \left[\sum_{([u,v],w) \in \pi_j} g_j(w)(v-u) + \varepsilon 2^{-j} \right] &< \\ \sum_{j=1}^N \left[\sum_{([u,v],w) \in \pi_j} \varepsilon(v-u) + \varepsilon 2^{-j} \right] &< \varepsilon(b-a+1). \end{aligned}$$

This shows that

$$0 \leq \int_a^b g_m(x) dx < \varepsilon(b-a+1)$$

for all $m \geq N$. The identity

$$\int_a^b f(x) dx - \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \lim_{n \rightarrow \infty} \int_a^b g_n(x) dx = 0.$$

follows.

Exercise 377, page 120

Just apply the theorems. We need, first, to determine that the interval of convergence of the integrated series

$$F(x) = \sum_{n=0}^{\infty} x^{n+1}/(n+1) = x + x^2/2 + x^3/3 + x^4/4 + \dots +$$

is $[-1, 1)$. Consequently, of the two integrals here, only one exists. Note that

$$F(0) - F(-1) = -F(-1) = 1 - 1/2 + 1/3 - 1/4 + 1/5 - 1/6 + \dots$$

is a convergent alternating series and provides the value of the integral

$$\int_{-1}^0 \left(\sum_{n=0}^{\infty} x^n \right) dx.$$

Note that the interval of convergence of the original series is $(-1, 1)$ but that is not what we need to know. We needed very much to know what the interval of convergence of the *integrated series* was.

Exercise 378, page 120

The formula

$$1 + x + x^2 + x^3 + x^4 + \dots = \frac{1}{1-x} \quad (-1 < x < 1)$$

is just the elementary formula for the sum of a geometric series. Thus we do not need to use series methods to solve the problem; we just need to integrate the function

$$f(x) = \frac{1}{1-x}.$$

We happen to know that

$$\int \frac{1}{1-x} dx = -\log(1-x) + C$$

on $(-\infty, 1)$ so this integral is easy to work with without resorting to series methods. The integral

$$\int_{-1}^0 \frac{1}{1-x} dx = -\log(1-0) - (-\log(1-(-1))) = \log 2.$$

For the first exercise you should have found a series that we now know adds up to $\log 2$.

Exercise 379, page 120

No, you are wrong. And don't call me 'Shirley.'

The condition we need concerns the integrated series, not the original series. The integrated series is

$$F(x) = x + x^2/2 + x^3/3 + x^4/4 + \dots +$$

and, while this diverges at $x = 1$, it converges at $x = -1$ since

$$1 - 1/2 + 1/3 - 1/4 + 1/5 - 1/6 + \dots$$

is an alternating harmonic series, known to be convergent by the convergent alternating series test. Theorem 3.46 then guarantees that the integral exists on $[-1, 0]$ and predicts that it might not exist on $[0, 1]$.

The mistake here can also be explained by the nature of the calculus integral. Remember that in order for a function to be integrable on an interval $[a, b]$ it does not have to be defined at the endpoints or even bounded near them. The careless student is fussing too much about the function being integrated and not paying close enough attention to the integrated series. We know that $F(x)$ is an antiderivative for f on $(-1, 1)$ so the only extra fact that we need for the integral $\int_{-1}^0 f(x) dx$ is that F is continuous on $[-1, 0]$. It is.

Exercise 380, page 120

Yes. Inside the interval $(-R, R)$ this formula must be valid.

Exercise 381, page 121

Yes. If we are sure that the closed, bounded interval $[a, b]$ is inside the interval of convergence (i.e., either $(-R, R)$ or $(-R, R]$ or $[-R, R)$ or $[-R, R]$) then this formula must be valid.

Exercise 382, page 121

Both the series

$$f(x) = 1 + 2x + 3x^2 + 4x^3 + \dots$$

and the formally integrated series

$$F(x) = x + x^2 + x^3 + \dots$$

have a radius of convergence 1 and an interval of convergence exactly equal to $(-1, 1)$. Theorem 3.46 assures us, only, that F is an indefinite integral for f on $(-1, 1)$.

But

$$F(x) = x + x^2 + x^3 + \cdots = \frac{x}{1-x} \quad (-1 < x < 1).$$

If we define

$$G(x) = \frac{x}{1-x} \quad (-1 \leq x < 1).$$

then G is continuous on $[-1, 0]$ and $G'(x) = F'(x) = f(x)$ on $(-1, 1)$. Consequently

$$\int_{-1}^0 f(x) dx = G(0) - G(-1) = 1/2.$$

We were not able to write

$$\int_{-1}^0 f(x) dx = F(0) - F(-1) = -1 + 1 - 1 + 1 - 1 + 1 - \cdots$$

because $F(-1)$ is not defined (the series for F diverges at $x = -1$).

Since $G(x)$ is unbounded near $x = 1$ there is no hope of finding an integral for f on $[0, 1]$.

Exercise 383, page 121

Here $R = 0$. Show that

$$\lim_{k \rightarrow \infty} k^k r^k = 0$$

for every $r > 0$. Conclude that the series must diverge for every $x \neq 0$.

Exercise 384, page 121

Do $R = 0$, $R = \infty$, and $R = 1$. Then for any $0 < s < \infty$ take your power series for $R = 1$ and make a suitable change, replacing x by sx .

Exercise 388, page 122

This follows immediately from Exercise 387 without any further computation.

Exercise 389, page 123

This follows immediately from the inequalities

$$\liminf_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| \leq \liminf_{k \rightarrow \infty} \sqrt[k]{|a_k|} \leq \limsup_{k \rightarrow \infty} \sqrt[k]{|a_k|} \leq \limsup_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right|,$$

which can be established by comparing ratios and roots, together with Exercise 386.

Exercise 390, page 123**Exercise 391, page 123****Exercise 392, page 123****Exercise 393, page 123****Exercise 394, page 124****Exercise 395, page 124****Exercise 400, page 125**

If the series converges absolutely at an endpoint $\pm R$ of the interval of convergence then

$$|a_0| + |a_1|R + |a_2|R^2 + |a_3|R^3 + \dots$$

converges. For each x in the interval $[-R, R]$,

$$|a_k(x)x^k| \leq |a_k|R^k.$$

By the Weierstrass M-test the series converges uniformly on $[-R, R]$.

The conclusion is now that

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

is a uniformly convergent power series on the interval $[-R, R]$ and so f is continuous. We know that

$$f'(x) = a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + \dots$$

is convergent at least on $(-R, R)$ and that this is indeed the derivative of f there. It follows that f' is integrable on $[-R, R]$ and that f is an indefinite integral on that interval.

Exercise 401, page 125

For the proof we can assume that the series

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

has a radius of convergence 1 and that the series converges nonabsolutely at $x = 1$. We can assume that the interval of convergence is $(-1, 1]$. Any other case can be transformed into this case.

Set

$$s_n = a_0 + a_1 + a_2 + a_3 + \dots + a_{n-1}$$

and note that, by our hypothesis that the power series converges at $x = 1$, this is a convergent series. The sequence $b_k(x) = x^k$ is nonnegative and decreasing on the interval $[0, 1]$. one of the versions of Abel's theorem (Exercise 330) applies in exactly this situation and so we can claim that the series

$$\sum_{k=0}^{\infty} a_k b_k(x)$$

converges uniformly on $[0, 1]$. This is what we wanted.

The conclusion is now that

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

is a uniformly convergent power series on the interval $[0, 1]$ and so f is continuous. We know that

$$f'(x) = a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + \dots$$

is convergent at least on $(-1, 1)$ and that this is indeed the derivative of f there. It follows that f' is integrable on $[0, 1]$.

We already know that f' is integrable on any interval $[a, 0]$ for $-1 < a < 0$. Thus f' is integrable on any interval $[a, 1]$ for $-1 < a < 0$, and thus integrable on any interval $[a, b] \subset (-1, 1]$.

To finish let us remark on the transformations needed to justify the first paragraph. If

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

converges at $x = 1$ then

$$g(x) = a_0 - a_1x + a_2x^2 - a_3x^3 + \dots$$

converges at $x = -1$ and

$$h(x) = a_0 + R^{-1}a_1x + R^{-2}a_2x^2 + R^{-3}a_3x^3 + \dots$$

converges at $x = R$.

Exercise 403, page 125

Write out the Cauchy criterion for uniform convergence on $(-r, r)$ and deduce that the Cauchy criterion for uniform convergence on $[-r, r]$ must then also hold.

Exercise 406, page 126

The best that can be concluded is that *if there is any series representation* for f valid at least in some interval $(-r, r)$ for $r > 0$, then

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$$

must be that series. But it is possible that there simply is no power series representation of a function, even assuming that it is f is infinitely often differentiable at $x = 0$.

Exercise 407, page 129

Each of these steps, carried out, will lead to the conclusion that the area is expressible as an integral. The first step is the assumption that area is additive. The second step assumes that area can be estimated above and below in this way. The last two steps then follow mathematically from the first two.

The loosest version of this argument requires taking the concept for granted and simply assuming that an accumulation argument will work for it. Thus $A(x)$ accumulates all of the area of the region between a and x . Now add on a small bit more to get $A(x+h)$. The bit more that we have added on is close to $f(\xi) \times h$ for some [or any] choice of ξ inside $(x, x+h)$.

We “conclude” immediately that $\int_a^x f(t) dt$ expresses completely the measurement $A(x)$ that we require. You should be aware here of where you are making an *additive* assumption and where you are making an assumption of *continuity*.

Exercise 410, page 130

Well in fact you merely memorized that the area of a circle of radius R is πR^2 . Then the area of a half-circle (assuming that it has an area) would be half of that (assuming that areas add up). Notice that by basing area on integration theory we are on firmer ground for all such statements.

Exercise 411, page 130

The top half of the circle is the curve $y = \sqrt{r^2 - x^2}$ and the bottom half is $y = -\sqrt{r^2 - x^2}$ both on the interval $-r \leq x \leq r$. Just apply Definition 3.48 (and hope that you have the skills to determine exactly what the integral is).

Exercise 412, page 130

“The difficulty that occurs with this test of integrands is somewhat subtle. If a quantity Q is equal to the integral of a function f , then every upper sum of f is larger than Q and every lower sum of f is smaller than Q . On the other hand, even with some applications occurring at the most elementary level, it is not possible to know *a priori* that upper and lower sums bound Q . One knows this only after showing in some other way that the integral of f equals Q . Consider, for example, the area between the graphs of the functions $g(x) = 1 + x^2$ and $h(x) = 2x^2$ on $[0, 1]$. While for a small $\Delta x > 0$, the maximum of $g(x) - h(x)$ on $[0, \Delta x]$ occurs at 0, no rectangle of height 1 and width Δx contains the region between the graphs over $[0, \Delta x]$, so it is not clear *a priori* that $1 \cdot \Delta x$ is larger than the area of that region. Of course there are several methods to justify the integral needed here . . . , but even for this simple example the ‘universal’ method of upper and lower sums fails, and Bliss’s theorem also fails, as a test for the integrand.”

... from Peter A. Loeb, *A lost theorem of the calculus*, The Mathematical Intelligencer, Volume 24, Number 2 (June, 2002).

One can just ignore the difficulty and accept Definition 3.48 as a correct interpretation of area. Or, we could use Definition 3.47 and insist that areas can be added and subtracted. In that way

$$\int_0^1 [g(x) - h(x)] dx = \int_0^1 g(x) dx - \int_0^1 h(x) dx$$

gets around the problem, since both of these areas and integrals allow an interpretation using the method of exhaustion.

Yet again, we could consider, instead, adjusted Riemann sums

$$\sum_{i=1}^n [g(\xi_i) - h(\xi_i^*)](x_i - x_{i-1})$$

that also approximate the same integral $\int_0^1 [g(x) - h(x)] dx$. Then, judicious choices of ξ_i and ξ_i^* can be made to return to an argument that follows the principles of the method of exhaustion.

Exercise 413, page 131

The geometric series certainly sums to the value 1. Now use the definition of the integral $\int_1^\infty x^{-2} dx$ to compute its value.

Exercise 416, page 135

First note that

$$\max\{|p|, |q|\} \leq \sqrt{p^2 + q^2} \leq |p| + |q|$$

for all real numbers p and q . Consequently, if we make any choice of points

$$a = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = b,$$

the sum

$$\sum_{i=1}^n \sqrt{[F(t_i) - F(t_{i-1})]^2 + [G(t_i) - G(t_{i-1})]^2}$$

has, as an upper bound,

$$\sum_{i=1}^n |F(t_i) - F(t_{i-1})| + |G(t_i) - G(t_{i-1})| \leq V(F, [a, b]) + V(G, [a, b]).$$

Consequently, for the length L of the curve,

$$L \leq V(F, [a, b]) + V(G, [a, b]).$$

In the other direction

$$\sum_{i=1}^n |F(t_i) - F(t_{i-1})| \leq \sum_{i=1}^n \sqrt{[F(t_i) - F(t_{i-1})]^2 + [G(t_i) - G(t_{i-1})]^2} \leq L.$$

Thus $V(F, [a, b]) \leq L$. The inequality $V(G, [a, b]) \leq L$ is similarly proved.

Exercise 417, page 135

We know that $F'(t)$ and $|F'(t)|$ are integrable on $[a, b]$. We also know that

$$V(F, [a, b]) \leq \int_a^b |F'(t)| dt.$$

Consequently F has bounded variation on $[a, b]$. Similarly G has bounded variation on $[a, b]$. It follows from Exercise 416 that the curve is rectifiable.

Let $\varepsilon > 0$ and choose points

$$a = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = b$$

so that

$$L - \varepsilon < \sum_{i=1}^n \sqrt{[F(t_i) - F(t_{i-1})]^2 + [G(t_i) - G(t_{i-1})]^2} \leq L.$$

The sum increases if we add points, so we will add all points at which the derivatives $F'(t)$ or $G'(t)$ do not exist.

In between the points in the subdivision we can use the mean-value theorem to select

$$t_{i-1} < \tau_i < t_i \quad \text{and} \quad t_{i-1} < \tau_i^* < t_i$$

so that

$$[F(t_i) - F(t_{i-1})] = F'(\tau_i) \quad \text{and} \quad [G(t_i) - G(t_{i-1})] = G'(\tau_i^*).$$

Consequently

$$L - \varepsilon < \sum_{i=1}^n \sqrt{[F'(\tau_i)]^2 + [G'(\tau_i^*)]^2} \leq L.$$

But the sums

$$\sum_{i=1}^n \sqrt{[F'(\tau_i)]^2 + [G'(\tau_i^*)]^2}$$

are approximating sums for the integral

$$\int_a^b \sqrt{[F'(t)]^2 + [G'(t)]^2} dt.$$

Here we are applying Theorem 3.21 since we have selected two points τ_i and τ_i^* from each interval, rather than one point as the simplest version of approximating Riemann sums would demand. We easily check that the function

$$H(p, q) = \sqrt{p^2 + q^2} \leq |p| + |q|$$

satisfies the hypotheses of that theorem.

Exercise 418, page 135

Use the Darboux property of continuous functions.

Exercise 421, page 136

Translating from the language of curves to the language of functions and their graphs:

The length of the graph would be the least number L so that

$$\sum_{i=1}^n [(x_i - x_{i-1})]^2 + [f(x_i) - f(x_{i-1})]^2 \leq L$$

for all choices of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

This would be finite if and only if f has bounded variation on $[a, b]$ and would be smaller than $(b - a) + V(f, [a, b])$.

A formula for this length, in the case when f is continuously differentiable on (a, b) with a bounded derivative,

would be

$$L = \int_a^b \sqrt{1 + [f'(x)]^2} dx.$$

Exercise 422, page 136

The function is continuously differentiable, Lipschitz and so certainly of bounded variation. Hence the curve

$$x = t, y = f(t) \quad (0 \leq t \leq 2)$$

is rectifiable.

The formula

$$L = \int_0^2 \sqrt{1 + \frac{1}{4}(e^{2x} - 2 + e^{-2x})} dx$$

is immediate. Calculus students would be expected to have the necessary algebraic skills to continue. “Completing the square” will lead to an integral that can be done by hand.

Exercise 423, page 140

On the interval $[a, b]$ with no additional points inserted this is exactly the trapezoidal rule. The general formula just uses the same idea on each subinterval.

Exercise 424, page 140

We can assume that f is twice continuously differentiable on $[a, b]$ and then apply integration by parts [twice] to the integral

$$\int_a^b (x-a)(b-x) f''(x) dx.$$

One integration by parts will give

$$\int_a^b (x-a)(b-x) f''(x) dx = (x-a)(b-x) f'(x) \Big|_{x=a}^{x=b} - \int_a^b [a+b-2x] f'(x) dx$$

and a second integration by parts on this integral will give

$$[2x - (a + b)]f(x) \Big|_{x=a}^{x=b} - 2 \int_a^b f(x) dx = (b - a)(f(a) + f(b)) - 2 \int_a^b f(x) dx.$$

Exercise 425, page 140

Again we can assume that f is twice continuously differentiable on $[a, b]$. Then the preceding exercise supplies

$$\begin{aligned} \int_a^b f(x) dx - \frac{f(a) + f(b)}{2}(b - a) &= -\frac{1}{2} \int_a^b (x - a)(b - x) f''(x) dx \\ &= -f''(\xi) \int_a^b (x - a)(b - x) dx = -f''(\xi) \frac{(b - a)^3}{12}, \end{aligned}$$

making sure to apply the appropriate mean-value theorem for the integral above.

Exercise 426, page 140

Again we can assume that f is twice continuously differentiable on $[a, b]$. Then the preceding exercises supply

$$\int_a^b f(x) dx - \frac{f(a) + f(b)}{2}(b - a) = -\frac{1}{2} \int_a^b (x - a)(b - x) f''(x) dx.$$

Now just use the fact that

$$\max_{x \in [a, b]} (x - a)(b - x) = \frac{(b - a)^2}{4}$$

to estimate

$$\int_a^b (x - a)(b - x) |f''(x)| dx.$$

Exercise 427, page 141

The preceding exercises should help.

Exercise 428, page 141

This is from Edward Rozema, *Estimating the error in the trapezoidal rule*, The American Mathematical Monthly, Vol. 87 (2), (1980), pages 124–128.

The observation just uses the fact that the usual error is exactly equal to

$$\sum_{k=1}^n -\frac{(b-a)^3}{12n^3} f''(\xi_i).$$

where here we are required to take appropriate points ξ_i in each interval

$$[x_{i-1}, x_i] = \left[a + \frac{(i-1)(b-a)}{n}, a + \frac{i(b-a)}{n} \right] \quad (i = 1, 2, 3, \dots, n)$$

If we rewrite this sum in a more suggestive way the theorem is transparent. Just check that this is exactly the same sum:

$$-\frac{(b-a)^2}{12n^2} \sum_{k=1}^n f''(\xi_i)(x_i - x_{i-1}).$$

We recognize the sum as a Riemann sum for the integral $\int_a^b f''(x) dx$ and that integral can be evaluated as $f'(b) - f'(a)$. [For large enough n the sum is close to the integral; this is all that is intended here.]

Rozema goes on to note that, since we have an explicit (if approximate) error, we may as well use it. Thus an improved trapezoidal rule is

$$\int_a^b f(x) dx \approx T_n - \frac{(b-a)^2}{12n^2} [f'(b) - f'(a)]$$

and the error estimate when using the improvement can be shown to be

$$\frac{f'''(\xi)(b-a)^5}{720n^4}$$

which is rather better than the error for the original trapezoidal rule.

Exercise 429, page 142

We can see (since the correct value of the integral is provided) that $n = 1$ or $n = 2$ is nowhere large enough. A simple trial-and-error approach might work. Look for a large value of n , compute the trapezoidal rule approximation and see if

we are close enough. Apart from being tedious, this isn't much of a "method." For one thing we do not expect normally to be asked such a question when the value is already guaranteed. More importantly, even if we could determine that $n = 50,000$ is large enough, how would we know that larger values of n are equally accurate. The trapezoidal rule eventually converges to the correct value, but it does not (in general) work out that the values get closer and closer to the correct value.

In the case here the situation is really quite simpler. Since the function $f(x) = e^{x^2}$ is convex [sometimes called concave up] on the interval $[0, 1]$ the trapezoidal rule always overestimates the integral. Each successive application for larger n will get closer as it will be smaller. So you could solve the problem using trial-and-error in this way.

If you know how to program then this is reasonable. On the web you can also find Java Applets that will do the job for you. For example, at the time of writing, a nice one is here

[www.math.ucla.edu/~ronmiech/Java Applets/Riemann/index.html](http://www.math.ucla.edu/~ronmiech/Java%20Applets/Riemann/index.html)

that allows you to input

{f(x)= exp(x^2)}

and select the number of subdivisions. It is perhaps more instructive to do some experimental play with such applets than to spend an equal time with published calculus problems.

A more sensible method, which will be useful in more situations, is to use the published error estimate for the trapezoidal rule to find how large n must be so that the error is small enough to guarantee nine decimal place accuracy.

The second derivative of $f(x) = e^{x^2}$ is

$$f''(x) = 2e^{x^2} + 4x^2e^{x^2}.$$

A simple estimate on the interval $[0, 1]$ shows that $2 \leq f''(\xi) \leq 6e = 16.30969097$ for all $0 \leq \xi \leq 1$.

We know that the use of the trapezoidal rule at the n th stage produces an error

$$\text{error} = -\frac{1}{12n^2} f''(\xi),$$

where ξ is some number between 0 and 1.

Consequently if we want an error less than $10^{-9}/2$ [guaranteeing a nine decimal accuracy] we could require

$$\frac{1}{12n^2} f''(\xi) \leq \frac{1}{12n^2} (16.30969097) < 10^{-9}/2.$$

So

$$n^2 > \frac{1}{12}(16.30969097)(2 \times 10^9)$$

or $n > 52138$ will do the trick. Evidently a trial-and-error approach might have been somewhat lengthy. Notice that this method, using the crude error estimate for the trapezoidal rule, guarantees that for all $n > 52138$ the answer provided by that rule will be correct to a nine decimal accuracy. It does not at all say that we *must* use n this large. Smaller n will doubtless suffice too, but we would have to use a different method to find them.

What we could do is use a lower estimate on the error. We have

$$\text{error} = -\frac{1}{12n^2} f''(\xi) \geq -\frac{2}{12n^2},$$

where ξ is some number between 0 and 1. Thus we could look for values of n for which

$$-\frac{2}{12n^2} > -10^{-9}$$

which occurs for $n^2 < \frac{1}{6}10^9$, or $n < 12909.9$. Thus, before the step $n = 12,909$ there must be an error in the trapezoidal rule which affects at least the ninth decimal place.

Exercise 430, page 144

To show that $\int_0^\infty x^n e^{-x} dx = n!$ first find a recursion formula for

$$I_n = \int_0^\infty x^n e^{-x} dx \quad (n = 0, 1, 2, 3, \dots)$$

by integration by parts. A direct computation shows that $I_0 = 1$ and an integration by parts shows that $I_n = nI_{n-1}$. It follows, by induction, that $I_n = n!$.

In fact Maple is entirely capable of finding the answer to this too. Input the same command:

```
> int(x^n* exp(-x), x=0..infinity );
memory used=3.8MB, alloc=3.1MB, time=0.38
```

GAMMA(n + 1)

The Gamma function is defined as $\Gamma(n+1) = n!$ at integers, but is defined at nonintegers too.

Exercise 432, page 145

This is called the *Cauchy-Schwarz inequality* and is the analog for integrals of that same inequality in elementary courses. It can be proved the same way and does not involve any deep properties of integrals.

Exercise 433, page 145

For example, prove the following:

1. $\log 1 = 0$.
2. $\log x < \log y$ if $0 < x < y$.
3. $\lim_{x \rightarrow \infty} \log x = \infty$ and $\lim_{x \rightarrow 0^+} \log x = 0$.
4. the domain and range are both $(0, \infty)$.
5. $\log xy = \log x + \log y$ if $0 < x, y$.
6. $\log x/y = \log x - \log y$ if $0 < x, y$.
7. $\log x^r = r \log x$ for $x > 0$ and $r = 1, 2, 3, \dots$
8. $\log e = 1$ where $e = \lim_{n \rightarrow \infty} (1 + 1/n)^n$.
9. $\frac{d}{dx} \log x = 1/x$ for all $x > 0$.
10. $\log 2 = 0.69\dots$
11. $\log(1+x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \dots$ for $-1 < x < 1$.

Exercise 436, page 148

Take any sequence. It must contain every element of $\emptyset$ since there is nothing to check.

Exercise 437, page 148

If the finite set is $\{c_1, c_2, c_3, \dots, c_m\}$ then the sequence

$$c_1, c_2, c_3, \dots, c_m, c_m, c_m, c_m, \dots$$

contains every element of the set.

Exercise 438, page 148

If the sequence

$$c_1, c_2, c_3, \dots, c_m, \dots$$

contains every element of some set it must certainly contain every element of any subset of that set.

Exercise 439, page 148

The set of natural numbers is already arranged into a list in its natural order. The set of integers (including 0 and the negative integers) is not usually presented in the form of a list but can easily be so presented, as the following scheme suggests:

$$0, 1, -1, 2, -2, 3, -3, 4, -4, 5, -5, 6, -6, 7, -7, \dots$$

Exercise 440, page 148

The rational numbers can also be listed but this is quite remarkable, for at first sight no reasonable way of ordering them into a sequence seems likely to be possible. The usual order of the rationals in the reals is of little help.

To find such a scheme define the “rank” of a rational number m/n in its lowest terms (with $n \geq 1$) to be $|m| + n$. Now begin making a finite list of all the rational numbers at each rank; list these from smallest to largest. For example, at rank 1 we would have only the rational number $0/1$. At rank 2 we would have only the rational numbers $-1/1, 1/1$. At rank 3 we would have only the rational numbers $-2/1, -1/2, 1/2, 2/1$. Carry on in this fashion through all the ranks. Now construct the final list by concatenating these shorter lists in order of the ranks:

$$0/1, -1/1, 1/1, -2/1, -1/2, 1/2, 2/1, \dots$$

This sequence will include every rational number.

Exercise 441, page 148

If the sequence

$$c_1, c_2, c_3, \dots, c_m, \dots$$

contains every element of a set A and the sequence

$$d_1, d_2, d_3, \dots, d_m, \dots$$

contains every element of a set B , then the combined sequence

$$c_1, d_1, c_2, d_2, c_3, d_3, \dots, c_m, d_m, \dots$$

contains every element of the union $A \cup B$.

By induction, then the union of any finite number of countable sets is countable. That is not so remarkable in view of the next exercise.

Exercise 442, page 148

We show that the following property holds for countable sets: If

$$S_1, S_2, S_3, \dots$$

is a sequence of countable sets of real numbers, then the set S formed by taking all elements that belong to at least one of the sets S_i is also a countable set.

We can consider that the elements of each of the sets S_i can be listed, say,

$$S_1 = \{x_{11}, x_{12}, x_{13}, x_{14}, \dots\}$$

$$S_2 = \{x_{21}, x_{22}, x_{23}, x_{24}, \dots\}$$

$$S_3 = \{x_{31}, x_{32}, x_{33}, x_{34}, \dots\}$$

$$S_4 = \{x_{41}, x_{42}, x_{43}, x_{44}, \dots\}$$

and so on. Now try to think of a way of listing all of these items, that is, making one big list that contains them all.

Describe in a systematic way a sequence that starts like this:

$$x_{11}, x_{12}, x_{21}, x_{13}, x_{22}, x_{23}, x_{14}, x_{23}, x_{32}, x_{41}, \dots$$

Exercise 443, page 148

First of all establish that for such a function and at every point $a < x \leq b$, the one-sided limit $F(x-) = \lim_{x \rightarrow x-} F(x)$ exists and that, at every point $a \leq x < b$, the one-sided limit $F(x+) = \lim_{x \rightarrow x+} F(x)$ exists. Note that, again because the function is monotonic, nondecreasing,

$$F(x-) \leq F(x) \leq F(x+)$$

at all $a < x < b$. Consequently F is continuous at a point x if and only if the one-sided limits at that point have the same value as $F(x)$.

For each integer n let C_n be the set of points x such that $F(x+) - F(x-) > 1/n$. Because F is nondecreasing, and because $F(b) - F(a)$ is finite there can only be finitely many points in any set C_n . To see this take, if possible, any points $a < c_1 < c_2 < \dots < c_p < b$ from the set C_n and select points

$$a = x_0 < c_1 < x_1 < c_2 < \dots < x_{p-1} < c_p < x_p = b.$$

Then

$$F(x_{i-1}) \leq F(c_i-) \leq F(c_i+) \leq F(x_i)$$

and so

$$p/n < \sum_{i=1}^p [F(c_i+) - F(c_i-)] \leq \sum_{i=1}^p [F(x_i) - F(x_{i-1})] = F(b) - F(a).$$

Thus the number of points in C_n cannot be larger than $n(F(b) - F(a))$.

The total set of points of discontinuity includes all the finite sets C_n together with (possibly) the points a and b . This set must be countable.

Exercise 445, page 149

If (a, b) is countable then find a function $f : (a, b) \rightarrow (0, 1)$ one-to-one onto and consider the sequence $f(s_n)$, where $\{s_n\}$ is a sequence that is claimed to have all of (a, b) as its range.

The simplest such function is, perhaps, $f(t) = (t - a)/(b - a)$. The same function shows that $[a, b]$ is countable if and only if $[0, 1]$ is countable. But if $[0, 1]$ is countable so is its subset $(0, 1)$. Indeed, if there exists a countable interval, then all intervals, open or closed, bounded or unbounded must be countable too.

Exercise 446, page 149

Recall that

1. Every number has a decimal expansion.
2. The decimal expansion is unique except in the case of expansions that terminate in a string of zeros or nines [e.g., $1/2 = 0.5000000 \dots = .49999999 \dots$], thus if a and b are numbers such that in the n th decimal place one has a 5 (or a 6) and the other does not then either $a \neq b$, or perhaps one ends in a string of zeros and the other in a string of nines.
3. Every string of 5's and 6's defines a real number with that decimal expansion.

We suppose that the theorem is false and that there is a sequence $\{s_n\}$ so that every number in the interval $(0, 1)$ appears at least once in the sequence. We may assume that all of the numbers of the sequence are in the interval $(0, 1)$ [otherwise remove them].

We obtain a contradiction by showing that this cannot be so. We shall use the sequence $\{s_n\}$ to find a number c in the interval $(0, 1)$ so that $s_n \neq c$ for all n .

Each of the points $s_1, s_2, s_3 \dots$ in our sequence is a number between 0 and 1 and so can be written as a decimal fraction. If we write this sequence out in decimal notation it might look like

$$s_1 = 0.x_{11}x_{12}x_{13}x_{14}x_{15}x_{16} \dots$$

$$s_2 = 0.x_{21}x_{22}x_{23}x_{24}x_{25}x_{26} \dots$$

$$s_3 = 0.x_{31}x_{32}x_{33}x_{34}x_{35}x_{36} \dots$$

etc. Now it is easy to find a number that is not in the list. Construct

$$c = 0.c_1c_2c_3c_4c_5c_6\dots$$

by choosing c_i to be either 5 or 6 whichever is different from x_{ii} . This number cannot be equal to any of the listed numbers $s_1, s_2, s_3 \dots$ since c and s_i differ in the i th position of their decimal expansions. This gives us our contradiction and so proves the theorem.

Exercise 447, page 149

Well, you could But you are missing the point of a proof by contradiction. To prove the theorem, we suppose that it fails and then obtain a contradiction from that assumption. Here we are supposing that we have succeeded in finding a listing of all the numbers from the interval $(0, 1)$. We construct a number that is not in the list and conclude that our assumption [that we have succeeded in finding a listing] is simply not valid.

Exercise 448, page 149

We suppose that the theorem is false and that there is a sequence $\{s_n\}$ so that every number in the interval (a, b) appears at least once in the sequence.

We obtain a contradiction by showing that this cannot be so. We shall use the sequence $\{s_n\}$ to find a number c in the interval (a, b) so that $s_n \neq c$ for all n .

Choose a subinterval $[c_1, d_1] \subset (a, b)$ that does not contain the first element s_1 of the sequence. Then choose a subinterval $[c_2, d_2] \subset [c_1, d_1]$ that does not contain the second element s_2 of the sequence. Continue inductively in this manner to produce a nested sequence of closed bounded intervals. There is at least one point c that belongs to each of these intervals and yet that point cannot appear in the sequence $\{s_n\}$.

Exercise 449, page 149

Find a way of ranking the algebraic numbers in the same way that the rational numbers were ranked in Exercise 440.

Try this for a rank: take the smallest number

$$n + |a_n| + |a_{n-1}| + \dots + |a_1| + |a_0|$$

as the rank of an algebraic number if it satisfies the equation

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0.$$

Now verify that there are only finitely many algebraic numbers at any rank. The union of the set of algebraic numbers at all the different ranks must then be countable.

Exercise 450, page 149

Every interval must contain infinitely many transcendental numbers otherwise that interval must be countable. The interval would then be countable itself, since it must then be contained in the union of the set of algebraic numbers [which is countable] and the set of transcendental numbers [which we imagine is countable]. In fact, then, the set of transcendental numbers in any interval must be uncountable.

Exercise 451, page 151

Exercise 166 can be used.

Exercise 452, page 151

The derivative of F exists at all points in $(0, 1)$ except at these corners $1/n$, $n = 2, 3, 4, 5, \dots$. If $a > 0$ then the interval $[a, 1]$ contains only finitely many corners. But the interval $(0, 1)$ contains countably many corners! Thus the calculus integral in both the finite set version and in the countable set version will provide

$$\int_a^b F'(x) dx = F(b) - F(a)$$

for all $0 < a < b \leq 1$. The claim that

$$\int_0^b F'(x) dx = F(b) - F(0)$$

for all $0 < b \leq 1$ can be made only for the new extended integral.

Exercise 453, page 151

The same proof that worked for the calculus integral will work here. We know that, for any bounded function f on an interval (a, b) , there is a uniformly continuous function on $[a, b]$ whose derivative is $f(x)$ at every point of continuity of f .

Exercise 455, page 152

Just kidding. But if some instructor has a need for such a text we could rewrite Chapters 2 and 3 without great difficulty to accommodate the more general integral. The discussion of countable sets in Chapter 4 moves to Chapter 1. The definition of the indefinite integral in Chapter 2 and the definite integral in Chapter 3 change to allow countable exceptional sets.

Most things can stay unchanged but one would have to try for better versions of many statements. Since this new integral is also merely a teaching integral we would need to strike some balance between finding the best version possible and simply presenting a workable theory that the students can eventually replace later on with the correct integration theory on the real line.

Exercise 456, page 152

We will leave the reader to search for an example of such a sequence. The exercise should leave you with the impression that the countable set version of the calculus integral is sufficiently general to integrate just about any example you could imagine creating. It is not hard to find a function that is not integrable by any reasonable method. But if it is possible (as this exercise demands) to write

$$f(x) = \sum_{k=1}^{\infty} g_k(x)$$

and if

$$\sum_{k=1}^{\infty} \left(\int_a^b g_k(x) dx \right)$$

converges then, certainly, f should be integrable. Any method that fails to handle f is inadequate.

With some work and luck you might consider the series

$$\sum_{k=1}^{\infty} \frac{a_k}{\sqrt{|x - r_k|}}$$

where $\sum_{k=1}^{\infty} a_k$ converges and $\{r_k\}$ is an enumeration of the rationals in $[0, 1]$. This is routinely handled by modern methods of integration but the Riemann integral and these two weak versions of the calculus integral collapse with such an example.

Exercise 457, page 153

Start with a set N that contains a single element c and show that that set has measure zero according to the definition. Let $\varepsilon > 0$ and choose $\delta(c) = \varepsilon/2$. Then if a subpartition

$$\{([c_i, d_i], c) : i = 1, 2\}$$

is given so that

$$0 < d_i - c_i < \delta(c) \quad (i = 1, 2)$$

then

$$\sum_{i=1}^2 (d_i - c_i) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Note that we have used only two elements in the subpartition since we cannot have more intervals in a subpartition with one associated point c .

Now consider a set $N = \{c_1, c_2, c_3, \dots, c_n\}$ that contains a finite number of elements. We show that that set has measure zero according to the definition. Let $\varepsilon > 0$ and choose $\delta(c_i) = \varepsilon/(2n)$ for each $i = 1, 2, 3, \dots, n$. Use the same argument but now with a few more items to keep track of.

Exercise 458, page 153

Now consider a countable set $N = \{c_1, c_2, c_3, \dots\}$ that contains an finite number of elements. We show that that set has measure zero according to the definition. Let $\varepsilon > 0$ and choose $\delta(c_k) = \varepsilon 2^{-k-1}$ for each $i = 1, 2, 3, \dots, n$. Use the same argument as in the preceding exercise but now with a quite a few more items to keep track of.

Suppose that we now have a subpartition

$$\{([c_i, d_i], \xi_i) : i = 1, 2, \dots, n\}$$

with each $\xi_i = c_k \in N$ for some k , and so that

$$0 < d_i - c_i < \delta(\xi_i) \quad (i = 1, 2, \dots, n).$$

Then to estimate the sum

$$\sum_{i=1}^n (d_i - c_i)$$

just check the possibilities where $([c_i, d_i], \xi_i) = ([c_i, d_i], c_k)$ for some k . Each of these adds no more than $2\varepsilon 2^{-k-1}$ to the value of the sum. But

$$\sum_{k=1}^{\infty} \varepsilon 2^{-k} = \varepsilon.$$

Exercise 459, page 153

Prove this by contradiction. If an interval $[a, b]$ does indeed have measure zero then, for any $\varepsilon > 0$, and every point $\xi \in [a, b]$ we should be able to find a $\delta(\xi) > 0$ with the following property: whenever a subpartition

$$\{([c_i, d_i], \xi_i) : i = 1, 2, \dots, n\}$$

is given with each $\xi_i \in [a, b]$ and so that

$$0 < d_i - c_i < \delta(\xi_i) \quad (i = 1, 2, \dots, n)$$

then

$$\sum_{i=1}^n (d_i - c_i) < \varepsilon.$$

By the Cousin covering argument there is indeed such a partition

$$\{([c_i, d_i], \xi_i) : i = 1, 2, \dots, n\}$$

with this property that is itself a full partition of the interval $[a, b]$. For that partition

$$\sum_{i=1}^n (d_i - c_i) = b - a.$$

This is impossible.

Exercise 460, page 153

To easy for a hint.

Exercise 461, page 153

We know that subsets of sets of measure zero have themselves measure zero. Thus if N_1 and N_2 are the two sets of measure zero, write

$$N_1 \cup N_2 = N_1 \cup [N_2 \setminus N_1].$$

The sets on the right are disjoint sets of measure zero. So it is enough if we prove the statement, assuming always that the two sets are disjoint and have measure zero.

Let $\varepsilon > 0$. To every point $\xi \in N_1$ or $\xi \in N_2$, there is a $\delta(\xi) > 0$ with the following property: whenever a subpartition

$$\{([c_i, d_i], \xi_i) : i = 1, 2, \dots, n\}$$

is given with each $\xi_i \in N_1$ or else with $\xi_i \in N_2$ and so that

$$0 < d_i - c_i < \delta(\xi_i) \quad (i = 1, 2, \dots, n)$$

then

$$\sum_{i=1}^n (d_i - c_i) < \varepsilon/2.$$

Together that means that whenever a subpartition

$$\{([c_i, d_i], \xi_i) : i = 1, 2, \dots, n\}$$

is given with each $\xi_i \in N_1 \cup N_2$ and so that

$$0 < d_i - c_i < \delta(\xi_i) \quad (i = 1, 2, \dots, n)$$

then

$$\sum_{i=1}^n (d_i - c_i) < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

since we can easily split the last sum into two parts depending on whether the associated points ξ_i belong to N_1 or belong to N_2 .

Exercise 462, page 153

We repeat our argument for the two set case but taking a little extra care. We know that subsets of sets of measure zero have themselves measure zero. Thus if $N_1, N_2, N_3, \dots$ is a sequence of sets of measure zero, write

$$N_1 \cup N_2 \cup N_3 \cdots = N_1 \cup [N_2 \setminus N_1] \cup (N_3 \setminus [N_1 \cup N_2]) \cup \dots$$

The sets on the right are disjoint sets of measure zero. So it is enough if we prove the statement, assuming always that the sets in the sequence are disjoint and have measure zero.

Let $\varepsilon > 0$. To every point $\xi \in N_k$ there is a $\delta(\xi) > 0$ with the following property: whenever a subpartition

$$\{([c_i, d_i], \xi_i) : i = 1, 2, \dots, n\}$$

is given with each $\xi_i \in N_k$ and so that

$$0 < d_i - c_i < \delta(\xi_i) \quad (i = 1, 2, \dots, n)$$

then

$$\sum_{i=1}^n (d_i - c_i) < \varepsilon 2^{-k}.$$

Together that means that whenever a subpartition

$$\{([c_i, d_i], \xi_i) : i = 1, 2, \dots, n\}$$

is given with each $\xi_i \in N_1 \cup N_2 \cup N_3 \cup \dots$ and so that

$$0 < d_i - c_i < \delta(\xi_i) \quad (i = 1, 2, \dots, n)$$

then

$$\sum_{i=1}^n (d_i - c_i) < \sum_{k=1}^p \varepsilon 2^{-k} < \varepsilon,$$

since we can easily split the last sum into finitely many parts depending on whether the associated points ξ_i belong to N_1 , or N_2 , or $N_3, \dots$, or N_p for some (possibly large) p .

Exercise 463, page 153

Let $\varepsilon > 0$. Since the series $\sum_{k=1}^{\infty} (b_k - a_k)$ converges there must be an integer N such that

$$\sum_{k=N}^{\infty} (b_k - a_k) < \varepsilon$$

. Note that every point of E is contained in one of the intervals (a_k, b_k) for $k = N, N+1, N+2, \dots$. For each $x \in E$ select the first one of these intervals (a_k, b_k) that contains x . Choose $\delta(x) < (b_k - a_k)/2$. This defines $\delta(x)$ for all x in E .

Whenever a subpartition

$$\{([c_i, d_i], \xi_i) : i = 1, 2, \dots, n\}$$

is given with each $\xi_i \in E$ and so that

$$0 < d_i - c_i < \delta(\xi_i) \quad (i = 1, 2, \dots, n)$$

then note that the interval $[c_i, d_i]$ belongs to one at least of the intervals (a_k, b_k) . Hence the sum

$$\sum_{i=1}^n (d_i - c_i)$$

can be split into a finite number of subsums each adding up to no more than $(b_k - a_k)$ for some $k = N, N+1, N+2, \dots$.

. It follows that

$$\sum_{i=1}^n (d_i - c_i) < \varepsilon.$$

Exercise 464, page 155

From each of the four closed intervals that make up the set K_2 remove the middle third open interval. This will lead to

$$K_3 = \left[0, \frac{1}{27}\right] \cup \left[\frac{2}{27}, \frac{3}{27}\right] \cup \dots$$

There should be eight intervals in all at this stage.

Exercise 466, page 156

First note that G is an open dense set in $[0, 1]$. Write $G = \bigcup_{k=1}^{\infty} (a_k, b_k)$. (The component intervals (a_k, b_k) of G can be called the intervals *complementary* to K in $(0, 1)$. Each is a middle third of a component interval of some K_n .) Observe that no two of these component intervals can have a common endpoint. If, for example, $b_m = a_n$, then this point would be an isolated point of K , and K has no isolated points.

Next observe that for each integer k the points a_k and b_k are points of K . But there are other points of K as well. In fact, we shall see presently that K is uncountable. These other points are all limit points of the endpoints of the complementary intervals. The set of endpoints is countable, but the closure of this set is uncountable as we shall see. Thus, in the sense of cardinality, “most” points of the Cantor set are *not* endpoints of intervals complementary to K .

Show that the remaining set $K = [0, 1] \setminus G$ is closed and nowhere dense in $[0, 1]$. Show that K has no isolated points and is nonempty. Show that K is a nonempty, nowhere dense perfect subset of $[0, 1]$.

Now let

$$G = \bigcup_{n=1}^{\infty} G_n$$

and let

$$K = [0, 1] \setminus G = \bigcap_{n=1}^{\infty} K_n.$$

Then G is open and the set K (our Cantor set) is closed.

To see that K is nowhere dense, it is enough, since K is closed, to show that K contains no open intervals. Let J be an open interval in $[0, 1]$ and let λ be its length. Choose a natural number n such that $1/3^n < \lambda$. By property 5, each component of K_n has length $1/3^n < \lambda$, and by property 2 the components of K_n are pairwise disjoint. Thus K_n cannot contain J , so neither can $K = \bigcap_{n=1}^{\infty} K_n$. We have shown that the closed set K contains no intervals and is therefore nowhere dense.

It remains to show that K has no isolated points. Let $x_0 \in K$. We show that x_0 is a limit point of K . To do this we show that for every $\varepsilon > 0$ there exists $x_1 \in K$ such that $0 < |x_1 - x_0| < \varepsilon$. Choose n such that $1/3^n < \varepsilon$. There is a component L of K_n that contains x_0 . This component is a closed interval of length $1/3^n < \varepsilon$. The set $K_{n+1} \cap L$ has two

components L_0 and L_1 , each of which contains points of K . The point x_0 is in one of the components, say L_0 . Let x_1 be any point of $K \cap L_1$. Then $0 < |x_0 - x_1| < \varepsilon$. This verifies that x_0 is a limit point of K . Thus K has no isolated points.

Exercise 467, page 156

Each component interval of the set G_n has length $1/3^n$; thus the sum of the lengths of these component intervals is

$$\frac{2^{n-1}}{3^n} = \frac{1}{2} \left(\frac{2}{3} \right)^n.$$

It follows that the lengths of all component intervals of G forms a geometric series with sum

$$\sum_{n=1}^{\infty} \frac{1}{2} \left(\frac{2}{3} \right)^n = 1.$$

(This also gives us a clue as to why K cannot contain an interval: After removing from the unit interval a sequence of pairwise disjoint intervals with length-sum one, no room exists for any intervals in the set K that remains.)

Exercise 468, page 156

Here is a hint that you can use to make into a proof. Let E be the set of all points in the Cantor set that are not endpoints of a complementary interval. Then the Cantor set is the union of E and a countable set. If E has measure zero, so too has the Cantor set.

Let $\varepsilon > 0$ and choose N so large that

$$\sum_{n=1}^N \frac{1}{2} \left(\frac{2}{3} \right)^n > 1 - \varepsilon.$$

i.e., so that

$$\sum_{n=N+1}^{\infty} \frac{1}{2} \left(\frac{2}{3} \right)^n < \varepsilon.$$

Here is how to define a $\delta(\xi)$ for every point in the set E . Just make sure that $\delta(\xi)$ is small enough that the open interval $(\xi - \delta(\xi), \xi, +\delta(\xi))$ does not contain any of the open intervals complementary to the Cantor set that are counted in the

sum

$$\sum_{n=1}^N \frac{1}{2} \left(\frac{2}{3} \right)^n > 1 - \epsilon.$$

Now check the definition to see that E satisfies the required condition to check that it is a set of measure zero. Using this δ guarantees that the intervals you will sum do not meet these open intervals that we have decided make up most of $[0, 1]$ (i.e., all but ϵ).

Exercise 472, page 157

This exercise shows that there is a purely arithmetical construction for the Cantor set. You will need some familiarity with ternary (base 3) arithmetic here.

Each $x \in [0, 1]$ can be expressed in base 3 as

$$x = .a_1a_2a_3\dots,$$

where $a_i = 0, 1$ or $2, i = 1, 2, 3, \dots$. Certain points have two representations, one ending with a string of zeros, the other in a string of twos. For example, $.1000\dots = .0222\dots$ both represent the number $1/3$ (base ten). Now, if $x \in (1/3, 2/3)$, $a_1 = 1$, thus each $x \in G_1$ must have '1' in the first position of its ternary expansion. Similarly, if

$$x \in G_2 = \left(\frac{1}{9}, \frac{2}{9} \right) \cup \left(\frac{7}{9}, \frac{8}{9} \right),$$

it must have a 1 in the second position of its ternary expansion (i.e., $a_2 = 1$). In general, each point in G_n must have $a_n = 1$. It follows that every point of $G = \bigcup_1^\infty G_n$ must have a 1 someplace in its ternary expansion.

Now endpoints of intervals complementary to K have two representations, one of which involves no 1's. The remaining points of K never fall in the middle third of a component of one of the sets K_n , and so have ternary expansions of the form

$$x = .a_1a_2\dots \quad a_i = 0 \text{ or } 2.$$

We can therefore describe K arithmetically as the set

$$\{x = .a_1a_2a_3\dots t \text{ (base three)} : a_i = 0 \text{ or } 2 \text{ for each } i \in N\}.$$

Exercise 473, page 157

In fact, K can be put into 1-1 correspondence with $[0,1]$: For each

$$x = .a_1a_2a_3 \dots \text{ (base 3), } a_i = 0, 2,$$

in the set K , let there correspond the number

$$y = .b_1b_2b_3 \dots \text{ (base 2), } b_i = a_i/2.$$

This provides a 1-1 correspondence between K (minus endpoints of complementary intervals) and $[0,1]$ (minus the countable set of numbers with two base 2 representations). By allowing these two countable sets to correspond to each other, we obtain a 1-1 correspondence between K and $[0,1]$.

Exercise 474, page 157

“When I was a freshman, a graduate student showed me the Cantor set, and remarked that although there were supposed to be points in the set other than the endpoints, he had never been able to find any. I regret to say that it was several years before I found any for myself.”

Ralph P. Boas, Jr, from *Lion Hunting & Other Mathematical Pursuits* (1995).

It is clear that there must be many irrational numbers in the Cantor ternary set, since that set is uncountable and the rationals are countable. Your job is to find just one.

Exercise 477, page 158

This is certainly true for some open sets, but not for all open sets. Consider $G = (0,1) \setminus C$ where C is the Cantor ternary set. The closure of G is all of the interval $[0,1]$ so that G and its closure do not differ by a countable set and contain many more points than the endpoints as the student falsely claims.

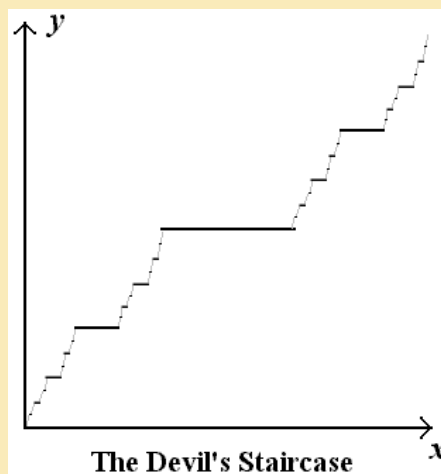


Figure 11.6: The Cantor function.

Exercise 481, page 161

See Donald R. Chalice, "A Characterization of the Cantor Function." Amer. Math. Monthly 98, 255–258, 1991 for a proof of the more difficult direction here, namely that the only monotone, nondecreasing function on $[0, 1]$ that has these three properties is the Cantor function. Figure 11.6 should be of assistance in seeing that each of the three properties holds. To verify them use the characterization of the function in the preceding exercise.

Exercise 482, page 162

There is nothing to prove. Write the two definitions and observe that they are identical.

Exercise 483, page 162

There is immediate. If the definition holds for the larger set then it holds without change for the smaller set.

Exercise 484, page 162

Let $\varepsilon > 0$. Then for every $x \in E_1$ there is a $\delta_1(x) > 0$

$$\sum_{i=1}^n |F(b_i) - F(a_i)| < \varepsilon/2$$

whenever a subpartition $\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$ is chosen for which

$$\xi_i \in E_1 \cap [a_i, b_i] \text{ and } b_i - a_i < \delta(\xi_i).$$

Similarly for every $x \in E_2$ there is a $\delta_2(x) > 0$

$$\sum_{i=1}^n |F(b_i) - F(a_i)| < \varepsilon$$

whenever a subpartition $\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$ is chosen for which

$$\xi_i \in E_2 \cap [a_i, b_i] \text{ and } b_i - a_i < \delta(\xi_i).$$

Take $\delta(x)$ in such a way, that if a point x happens to belong to both sets then $\delta(x)$ is the minimum of $\delta_1(x)$ and $\delta_2(x)$. For points that are not in both take $\delta(x)$ either $\delta_1(x)$ or $\delta_2(x)$.

Whenever a subpartition

$$\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$$

is chosen for which

$$\xi_i \in (E_1 \cup E_2) \cap [a_i, b_i] \text{ and } b_i - a_i < \delta(\xi_i)$$

the sum

$$\sum_{i=1}^n |F(b_i) - F(a_i)|$$

splits into two parts, depending on whether the ξ_i are in the first set E_1 or the second set E_2 . It follows that

$$\sum_{i=1}^n |F(b_i) - F(a_i)| < \varepsilon/2 + \varepsilon/2.$$

We have given all the details here since the next exercise requires the same logic but rather more detail.

Exercise 485, page 162

We can simplify the argument by supposing, without loss of generality, that the sets are disjoint. This can be arranged by using subsets of the E_j so that the union $E = \bigcup_{j=1}^{\infty} E_j$ is the same.

Let $\epsilon > 0$ and let $j = 1, 2, 3, \dots$. Then for every $x \in E_j$ there is a $\delta_j(x) > 0$

$$\sum_{i=1}^n |F(b_i) - F(a_i)| < \epsilon 2^{-j}$$

whenever a subpartition $\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$ is chosen for which

$$\xi_i \in E_j \cap [a_i, b_i] \text{ and } b_i - a_i < \delta_j(\xi_i).$$

Simply define $\delta(x) = \delta_j(x)$ if $x \in E_j$. Whenever a subpartition

$$\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$$

is chosen for which

$$\xi_i \in E \cap [a_i, b_i] \text{ and } b_i - a_i < \delta(\xi_i)$$

the sum

$$\sum_{i=1}^n |F(b_i) - F(a_i)|$$

splits into finitely many parts, depending on whether the ξ_i are in the first set E_1 , or the second set E_2 , or the third set E_3 , etc. It follows that

$$\sum_{i=1}^n |F(b_i) - F(a_i)| < \sum_{j=1}^{\infty} \epsilon 2^{-j} = \epsilon.$$

Exercise 486, page 163

If f is bounded on N then this is simple. Just use an upper bound, say $|f(x)| \leq M$ for $x \in N$ and note that

$$\sum_{i=1}^n |f(\xi_i)|(b_i - a_i) \leq M \sum_{i=1}^n (b_i - a_i).$$

If f is not bounded on N write, for every integer $j = 1, 2, 3, \dots$

$$N_j = \{x \in N : j - 1 \leq |f(x)| < j\}$$

and argue on each of these sets. Notice that we have zero variation on each set N_j since f is bounded on each set. The extension to the union of the sets $\{N_j\}$ is just a repetition of the details used in the proof of Exercise 485; just replace the sums

$$\sum_{i=1}^n |F(b_i) - F(a_i)|$$

by

$$\sum_{i=1}^n |f(\xi_i)|(b_i - a_i).$$

Exercise 487, page 163

This is particularly easy since

$$\sum_{i=1}^n |F(b_i) - F(a_i) - f(\xi_i)(b_i - a_i)| \leq \sum_{i=1}^n |F(b_i) - F(a_i)| + \sum_{i=1}^n |f(\xi_i)|(b_i - a_i).$$

Exercise 488, page 163

Select, for every $x \in E$, a $\delta(x) > 0$ so that

$$|F(v) - F(u) - f(x)(v - u)| < \frac{\varepsilon(v - u)}{b - a}$$

for all $0 < v - u < \delta(x)$ for which $u \leq x \leq v$. Then just check the inequality works since, if

$$\xi_i \in E \cap [a_i, b_i] \text{ and } b_i - a_i < \delta(\xi_i),$$

then

$$|F(b_i) - F(a_i) - f(\xi_i)(b_i - a_i)| < \frac{\varepsilon(b_i - a_i)}{b - a}.$$

Exercise 489, page 163

The Cantor function is, in fact, constant on each component of the open set complementary to the Cantor set in the interval $[0, 1]$. From that observation it is clear that the Cantor function has zero variation on each component interval of G . Then use Exercise 485.

Exercise 490, page 164

Let $\varepsilon > 0$. For every $x \in (a, b)$ there is a $\delta(x) > 0$ such that

$$\sum_{i=1}^n |F(b_i) - F(a_i)| < \varepsilon$$

whenever a subpartition $\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$ is chosen for which

$$\xi_i \in (a, b) \cap [a_i, b_i] \text{ and } b_i - a_i < \delta(\xi_i).$$

Consider any interval $[c, d] \subset (a, b)$. By the Cousin covering lemma there is a partition of the whole interval $[c, d]$, $\{([a_i, b_i], \xi_i) : i = 1, 2, \dots, n\}$, for which

$$\xi_i \in [a_i, b_i] \text{ and } b_i - a_i < \delta(\xi_i).$$

Consequently

$$|F(d) - F(c)| = \left| \sum_{i=1}^n F(b_i) - F(a_i) \right| \leq \sum_{i=1}^n |F(b_i) - F(a_i)| < \varepsilon.$$

This is true for any such interval and all positive ε . This is only possible if F is constant on (a, b) .

Exercise 491, page 164

We have already checked that the Cantor function has zero variation on the set complementary to the Cantor set in $[0, 1]$. This is because the Cantor function is constant on all of the component intervals. If the Cantor function also had zero variation on the Cantor set then we could conclude that it has zero variation on the entire interval $[0, 1]$. It would have to be constant.

Exercise 493, page 164

Just mimic (and simplify) the proof for Exercise 488.

Exercise 513, page 170

Note that you are required to prove that, if the conditions here hold, then indeed f would have to be integrable and

$$\int_a^b f(x) dx = F(b) - F(a).$$

Then you must supply a counterexample showing that not all integrable functions would necessarily have this property. Just review the section material on absolute continuity in Vitali's sense.

Exercise 514, page 170

If F is differentiable at all points of $[a, b]$ this is certainly a true statement. If we allow exceptional points then the hypotheses have to be adjusted.

Assume F is uniformly continuous on $[a, b]$ and differentiable at all but a countable set of points. Then this statement is true.

Assume F is Lipschitz on $[a, b]$ and differentiable at all but a set of points of measure zero. Then this statement is true. Remarkably enough this is true without assuming any differentiability. Lipschitz functions are always differentiable at all but a set of points of measure zero. But that observation belongs in Part Two of our text.

There are more conditions that you can assume to guarantee that

$$\int_a^b F'(x) dx = F(b) - F(a).$$

Exercise 518, page 173

Exercises 486, 487, and 488 contain all the pieces required for a very easy proof. Make sure to write

$$\int_{a_i}^{b_i} f(x) dx = F(b_i) - F(a_i)$$

using the indefinite integral F and to observe that only the first inequality of the theorem need be proved, since the second one follows immediately from the first.

Exercise 522, page 175

The proof is an exercise in derivatives taking care to handle the sets of measure zero. Use F and G for the indefinite integrals of f and g . Let N_0 be the set of points x in (a, b) where $f(x) \leq g(x)$ might fail. Suppose that $F'(x) = f(x)$ except on a set N_1 with N_1 measure zero and such that F has zero variation on N_1 . Suppose that $G'(x) = g(x)$ except on a set N_2 with N_2 measure zero and such that G has zero variation on N_2 .

Then $H = G - F$ has $H'(x) = g(x) - f(x) \geq 0$ except on the set $N_0 \cup N_1 \cup N_2$. This set is measure zero and, since F and G are absolutely continuous inside the interval, so too is H .

The proof then rests on the following fact which you should prove:

If H is uniformly continuous on $[a, b]$, absolutely continuous inside the interval, and if

$$\frac{d}{dx}H(x) \geq 0$$

for all points x in (a, b) except possibly points of a set of measure zero then $H(x)$ must be nondecreasing on $[a, b]$.

Finally then $H(a) \leq H(b)$ shows that $F(b) - F(a) \leq G(b) - G(a)$ and hence that

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

Exercise 523, page 175

Study the proof for Exercise 522 and just use those techniques here.

Exercise 524, page 176

In preparation ...

Exercise 525, page 176

Here is a version that is not particularly ambitious and is easy to prove. It is also sufficiently useful for most calculus classes. Suppose that F and G are uniformly continuous on $[a, b]$ and that each function is differentiable except at a countable number of points. Then the function $F(x)G'(x) + F'(x)G(x)$ is integrable on $[a, b]$ and

$$\int_a^b (F(x)G'(x) + F'(x)G(x)) \, dx = F(b)G(b) - F(a)G(a).$$

In particular $F(x)G'(x)$ is integrable on $[a, b]$ if and only if $F'(x)G(x)$ is integrable on $[a, b]$. In the event that either is integrable then the formula

$$\int_a^b F(x)G'(x) \, dx = F(b)G(b) - F(a)G(a) - \int_a^b F'(x)G(x) \, dx$$

must hold.

To prove it, just check that $H(x) = F(x)G(x)$ is uniformly continuous on $[a, b]$ and has a derivative at all but a countable number of points equal to the function $F(x)G'(x) + F'(x)G(x)$. But you can do better.

Exercise 526, page 177

Here are a number of versions that you might prove. Suppose G is uniformly continuous on $[a, b]$, and that F is uniformly continuous on an interval $[c, d]$ that includes every value of $G(x)$ for $a \leq x \leq b$. Suppose that each function is differentiable except at a countable number of points. Suppose that, for each $a \leq x \leq b$ the set

$$G^{-1}(G(x)) = \{t \in [a, b] : G(t) = G(x)\}$$

is at most countable. Then the function $F'(G(x))G'(x)$ is integrable on $[a, b]$ and

$$\int_a^b (F'(G(x))G'(x)) \, dx = F(G(b)) - F(G(a)).$$

To prove it, just check that $H(x) = F(G(x))$ is uniformly continuous on $[a, b]$ and has a derivative at all but a countable number of points equal to the function $F'(G(x))G'(x)$. Again you can do better. Try working with F and G as Lipschitz functions. Or take F everywhere differentiable and G as Lipschitz.

Exercise 527, page 177

There were an infinite number of points in the interval $[0, 1]$ at which we could not claim that

$$\frac{d}{dx}F(G(x)) = F'(G(x))G'(x).$$

But that set is countable and countable sets are no trouble to us now. So this function is integrable and the formula is valid.

Exercise 528, page 178

The proofs in Section 3.7.1 can be repeated with hardly any alterations. This is because both the calculus integral and the integral of this chapter can be given a pointwise approximation by Riemann sums. Just read through the proof and observe that the same arguments apply in this setting.

Exercise 530, page 179

The proofs in Section 3.7.1 can be repeated with hardly any alterations. This is because both the calculus integral and the integral of this chapter can be given a pointwise approximation by Riemann sums.

Exercise 531, page 179

Any constant function $F(x) = C$ will be, by definition, an indefinite integral for f .

Exercise 532, page 179

Any function $F(x)$ that is an indefinite integral for f will satisfy $F(d) - F(c) = 0$ for all $a \leq c < d \leq b$. Thus F is constant and $0 = F'(x) = f(x)$ for all x in the interval except possibly at points of a measure zero set.

Exercise 533, page 180

Any function $F(x)$ that is an indefinite integral for f will be monotonic, nondecreasing and satisfy $F(b) - F(a) = 0$. Thus F is constant and $0 = F'(x) = f(x)$ for all x in the interval except possibly at points of a measure zero set.

Exercise 535, page 191

Take any particular point x in E and check that $\beta(G)$ is full at that point x . Remember that, since G is open, there is a positive number δ_1 so that $(x - \delta_1, x + \delta_1) \subset G$. There is also a positive number δ_2 so that all pairs $([u, v], x)$ with $x \in [u, v]$ and $0 < v - u < \delta_2$ must belong to β .

Exercise 536, page 191

This is nearly identical to the preceding exercise, Exercise 535.

Exercise 549, page 193

This is a dual of the next exercise, Exercise 550.

Exercise 550, page 193

This is a dual of the preceding exercise, Exercise 549.

Exercise 553, page 193

For each x in E there would have to be at least one interval $(x, x + c)$ or $(x - c, x)$ that does not contain any points of the sequence.

Exercise 554, page 193

There would have to be at least one point x_0 in E at which β is not fine. That would mean that all intervals $(x, x + c)$ and $(x - c, x)$ contain infinitely many points of the sequence.

Exercise 555, page 193

For each x in E there would have to be at least one interval $(x - c, x + c)$ that does not contain any points of the sequence other than possibly x itself.

Exercise 556, page 194

There would have to be at least one point x_0 in E at which β is not full. That would mean that all intervals $(x_0, x_0 + c)$ or else all intervals $(x_0 - c, x_0)$ contain infinitely many points of the sequence.

Exercise 557, page 194

For each x in E every interval $(x, x + c)$ or else every interval $(x - c, x)$ contains infinitely many points of the sequence.

Exercise 558, page 194

There would have to be at least one point x_0 in E at which β is not fine. Thus some interval $(x_0, x_0 + c)$ or else some interval $(x_0 - c, x_0)$ contains no points of the sequence.

Exercise 559, page 194

For each x in E every interval $(x, x + c)$ and also every interval $(x - c, x)$ contains infinitely many points of the sequence.

Exercise 560, page 194

There would have to be at least one point x_0 in E at which β is not full. Thus some interval $(x_0, x_0 + c)$ or else some interval $(x_0 - c, x_0)$ contains no points of the sequence.

Exercise 567, page 199

The former represents a sum taken over *all* elements in the partition π while the latter sum contains only those elements (if any) $([u, v], w) \in \pi$ for which $[u, v]$ is a subinterval of $[c, d]$. It is the usual convention to consider that an empty sum has value zero.

Exercise 568, page 199

The former represents a sum taken over *all* elements in the partition π while the latter sum contains only those elements (if any) $([u, v], w) \in \pi$ for which w belongs to the set E . It is the usual convention to consider that an empty sum has value zero.

Exercise 573, page 202

It satisfies the definition easily, with $G = \emptyset$ in fact.

Exercise 574, page 202

If

$$E = \{x_1, x_2, \dots, x_N\}$$

and $\varepsilon > 0$, then the sequence of intervals

$$\left(x_i - \frac{\varepsilon}{2N}, x_i + \frac{\varepsilon}{2N}\right) \quad i = 1, 2, 3, \dots, N$$

covers the set E and the sum of all the lengths is ε . The union of these intervals is an open set G that contains E ; by the subadditivity property the Lebesgue measure $\lambda(G)$ is smaller than ε .

Exercise 575, page 202

If

$$E = \{x_1, x_2, \dots\}$$

and $\varepsilon > 0$, then the sequence of intervals

$$\left(x_i - \frac{\varepsilon}{2^{i+1}}, x_i + \frac{\varepsilon}{2^{i+1}}\right) \quad i = 1, 2, 3, \dots$$

covers the set E . Let G be the union of these intervals. Since

$$\sum_{k=1}^{\infty} 2 \left(\frac{\varepsilon}{2^{k+1}} \right) = \sum_{k=1}^{\infty} \varepsilon 2^{-k} = \varepsilon,$$

we conclude (from Lemma 5.8) that $\lambda(G) < \varepsilon$.

Exercise 576, page 202

Let $\varepsilon > 0$. Choose n so that $(2/3)^n < \varepsilon$. Then the n th stage intervals in the construction of the Cantor set give us 2^n closed intervals each of length $(1/3)^n$. This covers the Cantor set with 2^n closed intervals of total length $(2/3)^n$, which is less than ε . If the closed intervals trouble you (the definition requires open intervals), see Exercise 579 or argue as follows. Since $(2/3)^n < \varepsilon$ there is a positive number δ so that

$$(2/3)^n + \delta < \varepsilon.$$

Enlarge each of the closed intervals to form a slightly larger open interval, but change the length of each only enough so that the sum of the lengths of all the 2^n closed intervals does not increase by more than δ . The resulting collection of open intervals also covers the Cantor set, and the sum of the length of these intervals is less than ε . Thus the Cantor set has measure zero.

Exercise 578, page 203

Since E has measure zero, there is an open set G containing E for which $\lambda(G) < \varepsilon$. Let $\{(a_k, b_k)\}$ denote the component intervals of G . By the Heine-Borel theorem there is a finite N so that

$$\{(a_k, b_k) : k = 1, 2, \dots, N\}$$

covers the set E . Since

$$\sum_{k=1}^N (b_k - a_k) \leq \lambda(G) < \varepsilon.$$

the proof is complete.

Exercise 621, page 216

Don't forget to include the statement that F must be defined on an open interval that contains the point x_0 . You should verify that it means precisely that F is defined on an open set containing the point x_0 and is continuous at that point.

Exercise 636, page 222

Use Lemma 5.27.

Exercise 637, page 222

Let C the collection of points in (a, b) at which there is no derivative. This is countable and, since F is continuous, F has zero variation on C . Now take any measure zero set $N \subset (a, b)$. We know that F has zero variation on $C \cap N$ and, by Lemma 5.27, we know that F has zero variation on $N \setminus C$. It follows that F has zero variation on N .

Exercise 641, page 229

Our main tool, apart from ordinary computations, is the fact that monotonic functions are differentiable almost everywhere. This is proved in Theorem 5.28.

Let us simplify the proof by deciding that $F_n(a) = 0$ for all n , so that F and all functions F_n are nonnegative. We know from the Lebesgue differentiation theorem applied to all of these monotonic functions that, except for x in a set of measure zero, all of the derivatives, $F'(x)$ and $F'_n(x)$ exist. Thus it is only the identity for these values of x that we need to establish.

Note that

$$F(x) \geq \sum_{n=1}^m F_n(x)$$

for every integer m so that for almost every x ,

$$F'(x) \geq \sum_{n=1}^m F'_n(x)$$

and, consequently,

$$F'(x) \geq \sum_{n=1}^{\infty} F'_n(x). \quad (11.19)$$

To simplify we can assume that

$$F(b) - \sum_{n=1}^m F_n(b) \leq 2^{-m}.$$

If this were not the case then we could put parentheses in the series, group terms together, and relabel so that this would be the case. Consider the series

$$G(x) = \sum_{n=1}^{\infty} \left(F(x) - \sum_{k=1}^n F_k(x) \right).$$

Note that

$$0 \leq G(x) - \sum_{n=1}^m \left(F(x) - \sum_{k=1}^n F_k(x) \right) \leq \sum_{n=m+1}^{\infty} 2^{-n} = 2^{-m}.$$

Thus we see that G is also the sum of a series of functions.

A repeat of the argument we just gave to establish (11.19) will provide the analogous statement for this series:

$$0 \leq \sum_{n=1}^{\infty} \left(F'(x) - \sum_{k=1}^n F'_k(x) \right) \leq G'(x) \quad (11.20)$$

The function G has a finite derivative at almost every point. So in order for the inequality in (11.20) to hold for this series at a particular value of x the terms must tend to zero. Writing that out we now know that, for almost every x ,

$$\lim_{n \rightarrow \infty} \left(F'(x) - \sum_{k=1}^n F'_k(x) \right) = 0.$$

This is exactly the conclusion of the theorem.

Exercise 644, page 239

Make use in your proof of the fact that the intersection of two full covers, is again a full cover.

Exercise 647, page 239

Infinite values are allowed but we would have to avoid $\infty + (-\infty)$ or $-\infty + \infty$. This is simpler if you first check that a single value $f(b)$ is irrelevant to the computations so that you may assume that $f(b) = 0$. Then ensure that any partition π contained in your choice of β of the interval $[a, b]$, $[a, c]$ or $[b, c]$ would have to contain an element (I, b) .

Exercise 649, page 240

Check, first, that full covers do in fact contain endpointed partitions (as well as ordinary partitions). Then note that, if a partition π contains a pair $([u, v], w)$ for which $u < w < v$ that element can be replaced by the two items $([u, w], w)$ and $([w, v], w)$. That does not change the Riemann sums here because, for example,

$$f(w)[v - u] = f(w)[w - u] + f(w)[v - w].$$

Finally check that if β is a full cover there must be a smaller full cover $\beta' \subset \beta$ so that $([u, v], w) \in \beta'$ with $u < w < v$ if and only if both $([u, w], w)$ and $([w, v], w)$ are in β' .

Exercise 650, page 241

Use β to find estimates for the upper and lower integrals,

$$\overline{\int_a^b} f(x) dx - \underline{\int_a^b} f(x) dx < 2\varepsilon.$$

(Later we will show that this condition is, in fact, both necessary and sufficient.)

Exercise 657, page 244

Just check that the expression in Theorem 6.9

$$\left| \sum_{(I, w) \in \pi} \sum_{(I', w') \in \pi'} [f(w) - f(w')] \lambda(I \cap I') \right| < \varepsilon \quad (11.21)$$

is smaller than the expression

$$\sum_{(I, w) \in \pi} \sum_{(I', w') \in \pi'} |f(w) - f(w')| \lambda(I \cap I') < \varepsilon.$$

That would prove, using Theorem 6.9 that the integrability of f follows from the McShane criterion.

But it is also true that this expression

$$\left| \sum_{(I, w) \in \pi} \sum_{(I', w') \in \pi'} [|f(w)| - |f(w')|] \lambda(I \cap I') \right| < \varepsilon \quad (11.22)$$

is smaller than the expression

$$\sum_{(I,w) \in \pi} \sum_{(I',w') \in \pi'} |f(w) - f(w')| \lambda(I \cap I') < \varepsilon.$$

That would prove, again using Theorem 6.9 that the integrability of $|f|$ follows from the same McShane criterion.

Exercise 661, page 266

Check first that you need only prove the case where $E \subset (a, b)$. Then it is just a matter of looking carefully at the definitions of the two concepts.

Exercise 662, page 266

Use the subadditive property of open sets expressed in Lemma 5.8.

Exercise 668, page 282

Let A be the set of *all* points where f has a finite derivative. We know that each set of the form

$$\{x : \overline{D}f(x) > c\}$$

and

$$\{x : \underline{D}f(x) < c\}$$

is measurable. Thus the set A' of points at which f does not have a derivative (finite or infinite) can be expressed as the union of the family of sets

$$A_{pq} = \{x : \underline{D}f(x) < p < q < \overline{D}f(x)\}$$

for rational numbers p and q . It follows that A' is also measurable. Again the set A'' of points where $f'(x) = \pm\infty$ can be written as

$$A'' = \bigcap_{n=1}^{\infty} \{x : \overline{D}f(x) < -n\} \cup \bigcap_{n=1}^{\infty} \{x : \underline{D}f(x) > n\}.$$

Thus this set is measurable. But $A = A' \setminus A''$.

Exercise 675, page 291

Apply Fatou's lemma to the non-negative sequence given by $g f_n$.

Exercise 676, page 291

If f denotes the pointwise limit of the sequence, then f is also measurable and dominated by g , hence integrable. Furthermore,

$$|f(x) - f_n(x)| \leq 2g(x)$$

for all n and

$$\limsup_{n \rightarrow \infty} |f(x) - f_n(x)| = 0.$$

By the reverse Fatou lemma,

$$\limsup_{n \rightarrow \infty} \int_a^b |f(x) - f_n(x)| dx \leq \int_a^b \limsup_{n \rightarrow \infty} |f(x) - f_n(x)| dx = 0.$$

Using linearity and monotonicity of the integral,

$$\left| \int_a^b f(x) d\mu - \int_a^b f_n(x) dx \right| = \left| \int_a^b f(x) - f_n(x) dx \right| \leq \int_a^b |f(x) - f_n(x)| dx,$$

and the statement is proved.

Exercise 678, page 297

Let $\varepsilon > 0$ and suppose that $F'(x) = f(x)$ at every point. Define

$$\beta = \{(I, x) : |\Delta F(I) - f(x)\lambda(I)| < \varepsilon\lambda(I)\}.$$

Check that β is a full cover of $\mathbb{R}$ and that it satisfies (7.16).

Conversely suppose that β has been chosen to be a full cover of $\mathbb{R}$ that satisfies (7.16). Fix $x \in \mathbb{R}$ and determine $\delta > 0$ so that whenever (I, x) satisfies $x \in I$ and $\lambda(I) < \delta$ then necessarily $(I, x) \in \beta$. Note that if $([c, d], x)$ is any pair for which $c \leq x \leq d$ and $d - c < \delta$ then necessarily $([c, d], x)$ is in β and the set π containing only this one pair is itself also

a partition of $[c, d]$. Consequently, using (7.16),

$$|F(d) - F(c) - f(x)(d - c)| < \varepsilon(d - c).$$

But this verifies that $F'(x) = f(x)$.

Exercise 679, page 298

Suppose that $F'(x) = f(x)$ everywhere. Apply Lemma 678 to find a full cover β for which for every compact interval $[a, b]$ and every partition $\pi \subset \beta$ of $[a, b]$,

$$\sum_{(I,x) \in \pi} |\Delta F(I) - f(x)\lambda(I)| < \varepsilon\lambda([a, b])/2. \quad (11.23)$$

Let $\pi_1, \pi_2 \subset \beta$ be partitions of $[a, b]$. Apply (11.23) to each of them and add to obtain that

$$\left| \sum_{(I,z) \in \pi} f(z)\lambda(I) - \sum_{(I',z') \in \pi'} f(z')\lambda(I') \right| < \varepsilon\lambda([a, b]). \quad (11.24)$$

Now simply rearrange (11.24) to obtain (7.17). Conversely suppose that the statement (7.17) in the theorem holds for ε and β . This is a stronger statement than the Cauchy second criterion and so f is integrable on every compact interval. Thus there is a function F that will serve as the indefinite integral for f on any interval. From (7.17) we deduce that

$$\left| \Delta F([a, b]) - \sum_{(I,z) \in \pi} f(z)\lambda(I) \right| < 2\varepsilon\lambda([a, b]) \quad (11.25)$$

must be true for any partition π of $[a, b]$ from the cover β .

Fix $x \in \mathbb{R}$ and determine $\delta > 0$ so that whenever (I, x) satisfies $x \in I$ and $\lambda(I) < \delta$ then necessarily $(I, x) \in \beta$. Note that if $([c, d], x)$ is any pair for which $c \leq x \leq d$ and $d - c < \delta$ then necessarily $([c, d], x)$ is in β and the set π containing only this one pair is itself also a partition of $[c, d]$. Consequently, using (11.25),

$$|F(d) - F(c) - f(x)(d - c)| < 2\varepsilon(d - c).$$

But this verifies that $F'(x) = f(x)$.

Exercise 681, page 300

For any covering relation β it is clear that

$$V(r\lambda, \beta[E]) \leq V(f\lambda, \beta) \leq V(s\lambda, \beta)$$

and from this one can deduce that

$$r\lambda(E) = V^*(r\lambda, E) \leq V^*(f\lambda, E) \leq V^*(s\lambda, E) = s\lambda(E).$$

Note that it would also be true that

$$r\lambda(E) = V_*(r\lambda, E) \leq V_*(f\lambda, E) \leq V_*(s\lambda, E) = s\lambda(E).$$

Exercise 685, page 301

This exercise asserts that when $\lambda(E)$ is zero so too is $\int_E f(x) dx$. This is considered an absolute continuity condition. In Exercise 693 we consider a different version of absolute continuity asserting that if $\lambda(E)$ is “small” so too is $\int_E f(x) dx$.

Let $E_n = \{x \in E : 1/n < f(x)\}$. Check that

$$n\lambda(E_n) \leq \int_{E_n} f(x) dx \leq \int_E f(x) dx = 0.$$

Thus $\lambda(E_n) = 0$ for each n and so also if $E' = \{x \in E : f(x) \neq 0\}$ then

$$\lambda(E') \leq \sum_{n=1}^{\infty} \lambda(E_n) = 0.$$

Exercise 692, page 302

For illustrative purposes only we begin the proof with the bounded case. Suppose that $f(x) < N$ for all $x \in E$. Choose $\delta = \varepsilon/N$ and observe that, if $\lambda(G) < \delta$ then the inequalities in the measure estimates of an earlier exercise in this section provide

$$\int_{E \cap G} f(x) dx \leq N\lambda(G) < \varepsilon.$$

Thus the proof in the bounded case is trivial and does not require that f be measurable.

Exercise 693, page 302

This exercise asserts that when $\lambda(E)$ is small so too is $\int_E f(x) dx$. This is considered an absolute continuity condition. In Exercise 686 we consider a different version of absolute continuity asserting that if $\lambda(E)$ is zero so too is $\int_E f(x) dx$. Note that there are finiteness assumptions in this (stronger) version.

The argument in the preceding exercise suggests how to proceed. Let

$$A_n = \{x : n-1 \leq f(x) < n\}.$$

From the fact that f is measurable we can deduce that A_n is measurable. Thus we can select an open set G_n for which $B_n = A_n \setminus G_n$ is closed and $\lambda(G_n) < \varepsilon 2^{-n} n^{-1}$. That also requires

$$\begin{aligned} \int_{E \cap A_n} f(x) dx &\leq \int_{E \cap B_n} f(x) dx + \int_{E \cap A_n \cap G_n} f(x) dx \\ &\leq \int_{E \cap B_n} f(x) dx + \varepsilon 2^{-n}. \end{aligned}$$

Note that $\{B_n\}$ is a disjointed sequence of closed sets whose union B can be handled by the usual additive properties of measures over such sets. Thus

$$\begin{aligned} \sum_{n=1}^{\infty} \int_{E \cap A_n} f(x) dx &\leq \sum_{n=1}^{\infty} \int_{E \cap B_n} f(x) dx + \varepsilon \\ &= \int_{E \cap B} f(x) dx + \varepsilon \leq \int_E f(x) dx + \varepsilon < \infty. \end{aligned}$$

In particular there must be an integer N sufficiently large that

$$\sum_{n=N+1}^{\infty} \int_{E \cap A_n} f(x) dx < \varepsilon/2.$$

Choose $\delta = \varepsilon/(2N)$ and let G be any open set for which $\lambda(G) < \delta$. Since

$$E \cap G = \{x \in E \cap G : f(x) < N\} \cup \bigcup_{n=N+1}^{\infty} (G \cap E \cap A_n)$$

we have

$$\int_{E \cap G} f(x) dx \leq N\lambda(G) + \sum_{n=N+1}^{\infty} \int_{E \cap A_n} f(x) dx < \varepsilon.$$

Exercise 694, page 303

The proof repeats a number of techniques we have already seen in the proof of Theorem 693.

Each of the sets appearing in the statement of the theorem is measurable, because f is measurable. Select open sets G_{kr} so that $B_{rk} = A_{kr} \setminus G_{kr}$ is closed and so that

$$\lambda(G_{kr}) < \varepsilon 2^{-|k|-1} r^{-k}.$$

That also requires

$$\begin{aligned} \int_{E \cap A_{kr}} f(x) dx &\leq \int_{E \cap B_{kr}} f(x) dx + \int_{E \cap A_{kr} \cap G_{kr}} f(x) dx \\ &\leq \int_{E \cap B_{kr}} f(x) dx + \varepsilon 2^{-|k|-1}. \end{aligned}$$

Note that $\{B_{kr}\}$ is a disjointed sequence of closed sets whose union B_r can be handled by the usual additive properties of measures over such sets.

Now we compute:

$$\begin{aligned} \int_E f(x) dx &\leq \sum_{k=-\infty}^{\infty} \int_{E \cap A_{kr}} f(x) dx \\ &\leq \sum_{k=-\infty}^{\infty} \int_{E \cap B_{kr}} f(x) dx + \varepsilon \leq \sum_{k=-\infty}^{\infty} r^k \lambda(E \cap B_{kr}) + \varepsilon \\ &\leq r \sum_{k=-\infty}^{\infty} \int_{E \cap B_{kr}} f(x) dx + \varepsilon = r \int_{E \cap B_r} f(x) dx + \varepsilon \leq r \int_E f(x) dx + \varepsilon. \end{aligned}$$

Exercise 695, page 303

The first identity

$$\int_E f(x) dx = V^*(f\lambda, E)$$

is just our definition. Thus the intent of the exercise is to prove just that

$$V^*(f\lambda, E) = V_*(f\lambda, E).$$

Repeat Exercise 694 and this time deduce the related inequality

$$V_*(f\lambda, E) \leq \sum_{k=-\infty}^{\infty} r^k \lambda(E \cap A_{kr}) \leq r V_*(f\lambda, E).$$

Essentially this is accomplished because the Lebesgue measure can be estimated by either full covers or by fine covers (this is the Vitali covering theorem).

A comparison of the two inequalities shows that $V^*(f\lambda, E) = V_*(f\lambda, E)$.

Exercise 704, page 308

Define $F : \mathbb{R} \rightarrow \mathbb{R}$ by $F(0) = F(1/(2n-1)) = 0$ and $F(1/2n) = 1/n$ for all $n = 1, 2, 3, \dots$. Extend F to be linear on each of the intervals contiguous to these points where it has so far been defined. Show that F is absolutely continuous but that Vitali's condition does not hold on the interval $[0, 1]$.

Exercise 705, page 309

Let E be a null set. Write $E_n = E \cap (-n, n)$ for any integer n . We show that F has zero variation on E_n .

Using the δ, ε of the Vitali definition on the interval $[-n, n]$ cover E_n with a subsequence of open intervals $\{(c_i, d_i)\}$ with total length less than δ . Let β be the collection of all pairs $([u, v], w)$ for which $w \in E_n$ and $[u, v]$ is a subset of one at least of the open intervals $\{(c_i, d_i)\}$. This collection β is a full cover of E_n .

Let π be any subpartition contained in β . It must be the case, by the way that β has been constructed, that

$$\sum_{([u, v], w) \in \pi} (v - u) < \delta.$$

Consequently

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)| < \varepsilon.$$

From this it follows that F has zero variation on E_n . Since E is the union of the sequence of sets $\{E_n\}$ it follows too that F has zero variation on E .

Exercise 706, page 309

Establish that

$$|F(d) - F(c)| \leq M(d - c)$$

for some M and all $[c, d] \subset [a, b]$.

Exercise 708, page 309

Perhaps hard to spot. Note that the condition does not specify that the intervals should be nonoverlapping.

Exercise 712, page 311

Let M be an upper bound for the values of $|f(x)|$ in the interval. Let N be the measure zero set that allows us to say that f is continuous at every point in $[a, b] \setminus N$. We will assume that f is constant on $(-\infty, a]$ and on $[b, \infty)$. This just allows us to ignore what is happening outside of the interval $[a, b]$.

Let $\varepsilon > 0$ and define

$$\beta_1 = \{([x, y], z) : \omega f([x, y]) < \varepsilon/4(b - a).\}$$

Check, using the continuity of f , that β is a full cover of $\mathbb{R} \setminus N$. Verify that if (I, z) and (I', z') both belong to β then, either I and I' have no points in common or else $|f(z) - f(z')| < \varepsilon/(b - a)$.

Choose an open set G containing N with $\lambda(G) < \varepsilon/(2M)$. Let β_2 be the collection of all pairs $([u, v], w)$ for which $w \in N$, $u \leq w \leq v$, and $[u, v] \subset G$. It is easy to check that β_2 is a full cover of N . Thus $\beta = \beta_1 \cup \beta_2$ is a full cover of the whole real line.

Complete the proof by checking that

$$\sum_{(I,x) \in \pi} \sum_{(I',x') \in \pi'} |f(x) - f(x')| \lambda(I \cap I') < \varepsilon$$

for any pair of partitions π and π' of $[a, b]$. This just requires handling the pairs of items (I, x) and (I', x') differently depending on whether both came originally from β_1 or one of the pair is in β_2 . The first case we have already done in the preceding theorem. The second case should present no difficulties if the reader will remember the minor point that

$$|f(x) - f(x')| \leq 2M$$

in the sum.

Thus f satisfies McShane's criterion on $[a, b]$. It follows that f is Lebesgue integrable there.

Exercise 713, page 311

Let

$$E_n = \{x \in [a, b] : \omega_f(x) \geq 1/n\}.$$

This is a closed set containing only points of discontinuity of f . Indeed $\bigcup_{n=1}^{\infty} E_n$ is exactly the set of all points of discontinuity of f .

Fix n and $\varepsilon > 0$. Choose a partition π of $[a, b]$ so that

$$\sum_{(I,x) \in \pi} \omega_f(I) \lambda(I) < \varepsilon.$$

Let π' denote the subset of π containing just the elements $([x, y], z) \in \pi$ for which $(x, y) \cap E_n \neq \emptyset$. Note that each pair $([x, y], z)$ that belongs to π' will necessarily require

$$\omega_f([x, y]) \geq 1/n.$$

Also the collection of such intervals $[x, y]$ will cover all but finitely many points of E_n . Hence we compute that

$$\lambda(E_n) \leq \sum_{(I,x) \in \pi'} \lambda(I) \leq \sum_{(I,x) \in \pi'} n \omega_f(I) \lambda(I) < n\varepsilon.$$

This can happen for all ε only if $\lambda(E_n) = 0$ from which it follows that $\bigcup_{n=1}^{\infty} E_n$ (the set of points of discontinuity of f) is also of λ -measure zero. This proves (2).

Exercise 714, page 311

Finally let us assume (2) and show that this implies (1). Let $\eta > 0$ and choose M so that $|f(x)| < M$ for all x in $[a, b]$. Let

$$E_\eta = \{x \in [a, b] : \omega_f(x) \geq \eta\}.$$

This is a set of measure zero. There must be an open set G containing E_η so that $\lambda(G) < \eta/(2M)$. For each point x in $[a, b]$ that is not in G note that $\omega_f(x) < \eta$. Construct the covering relation

$$\beta_1 = \{(I, x) : x \in [a, b] \setminus G, x \in I, \text{ and } \omega_f(I) < \eta\}.$$

This is a full cover of $[a, b] \setminus G$. Construct the covering relation

$$\beta_2 = \{(I, x) : x \in G, x \in I, \text{ and } I \subset G\}.$$

Observe that $\beta = \beta_1 \cup \beta_2$ is a Cousin cover of $[a, b]$.

Let π be any partition of $[a, b]$ contained in β . Write $\pi' = \pi[G]$ and $\pi'' = \pi \setminus \pi[G]$. Then

$$\begin{aligned} \sum_{(I,x) \in \pi} \omega_f(I) \lambda(I) &= \sum_{(I,x) \in \pi'} \omega_f(I) \lambda(I) + \sum_{(I,x) \in \pi''} \omega_f(I) \lambda(I) \\ &\leq 2M \lambda(G) + \eta(b-a) < \eta(1+b-a). \end{aligned}$$

Since η is an arbitrary positive number we have deduced (1) from (2). In fact we have proved rather more, that the integral

$$\int_a^b \omega_f \cdot dx = 0$$

where that integral is interpreted as the usual limit of this kind of Riemann sums:

$$\sum_{(I,x) \in \pi} \omega_f(I) \lambda(I).$$

Exercise 715, page 311

We show first that a function f on an interval $[a, b]$ that is Riemann integrable [i.e., integrable but using uniformly full covers *must* be bounded and a.e. continuous. [This is an historically interesting fact, showing exactly the limitations of the Riemann integration theory.]

Use the oscillation $\omega_f(x)$ of a function f at a point x . (This value is positive if and only if f is discontinuous at x .) Check first the easy fact that f must be bounded. Fix $e > 0$ and consider the set $N(e)$ of points x such that the oscillation of f at x is greater than e ; that is, so that

$$\omega_f(x) > e.$$

Any interval (c, d) that contains a point $x \in N(e)$ will certainly have

$$\omega f([c, d]) \geq e.$$

Let $\varepsilon > 0$ and use Exercise 655 to find

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

such that

$$\sum_{k=1}^n \omega f([x_{k-1}, x_k]) (x_k - x_{k-1}) < \varepsilon e / 2.$$

Select just those intervals that contain a point from $N(e)$ in their interior. The total length of these intervals cannot exceed $(\varepsilon e) / (2e)$ since $\omega f([x_{k-1}, x_k]) \geq e$ for each interval $[x_{k-1}, x_k]$.

This covers the set $N(e)$ by a sequence of intervals $[x_{k-1}, x_k]$ of total length less than $\varepsilon / 2$, except that possibly we have missed a point x_i that happens to be in $N(e)$. In any case, argue that $N(e)$ has measure zero. But the set of points of discontinuity of f is the union of the sets $N(1)$, $N(1/2)$, $N(1/4)$, $N(1/8)$, $\dots$

Exercise 721, page 316

Take as a full cover β the collection of pairs $([u, v], w)$ for which $w \in [u, v]$ but $[u, v]$ never overlaps both of the intervals $[0, 1/2]$ or $[1/2, 1]$ unless $w = 1/2$. Then all partitions π of $[a, b]$ from β can be split neatly at the point $1/2$.

Exercise 722, page 316

Take as a full cover β the collection of pairs $([u, v], w)$ for which $w \in [u, v]$ but $[u, v]$ never overlaps two of the intervals $[\xi_{i-1}, \xi_i]$ unless w is one of the points $\{\xi_i\}$. Then all partitions π of $[a, b]$ from β can be split neatly at the points ξ_i .

Exercise 723, page 316

Both integrals exist but have different values, which you can check. If you were schooled in the Riemann-Stieltjes integral then you might recall this example was used to illustrate *non-existence* of the Riemann-Stieltjes integral. These differences in the two theories are mostly irrelevant since most applications will assume that one function is continuous and the other has bounded variation.

Exercise 724, page 317

Warning: If you were schooled in the Riemann-Stieltjes integral before learning *this* Stieltjes integral you may think not. Otherwise just check that the existence of the integral (finitely that is) on $[a, b]$ and $[b, c]$ is equivalent to the existence of the integral on $[a, c]$.

Exercise 731, page 317

Hint: $|dG(x)|$ is subadditive whereas $dG(x)$ is additive.

Exercise 749, page 340

We can simplify the argument and assume that F is defined on the whole real line. We wish to show that

1. $\lambda_F(\emptyset) = 0$.
2. For any sequence of sets $E, E_1, E_2, E_3, \dots$ for which $E \subset \bigcup_{n=1}^{\infty} E_n$ the inequality

$$\lambda_F(E) \leq \sum_{n=1}^{\infty} \lambda_F(E_n)$$

must hold.

This result is often described by the following language that splits the property (2) in two parts:

Subadditivity: $\lambda_F\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} \lambda_F(E_n).$

Monotonicity: $\lambda_F(A) \leq \lambda_F(B)$ if $A \subset B$.

The monotonicity is obvious. This allows us to prove the additivity assertion above just in the special case that the sets $\{E_n\}$ are pairwise disjoint so that $E = \bigcup_{n=1}^{\infty} E_n$ is now a disjoint union. If $\lambda_F(E_n) = \infty$ for any integer n there is nothing to prove so we may suppose all of these are finite.

Let $\varepsilon > 0$. For each integer n choose a full cover β_n of E_n so that

$$\sup_{\pi \subset \beta_n} \sum_{([u,v],w) \in \pi} |F(v) - F(u)| < \lambda_F(E_n) + \varepsilon 2^{-n}.$$

Then write

$$\beta = \bigcup_{n=1}^{\infty} \beta_n[E_n].$$

This is a full cover of E and consequently

$$\lambda_F(E) \leq \sup_{\pi \subset \beta} \sum_{([u,v],w) \in \pi} |F(v) - F(u)|.$$

Take any subpartition $\pi \subset \beta$ and observe that

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)| \leq \sum_{n=1}^{\infty} \sum_{([u,v],w) \in \pi[E_n]} |F(v) - F(u)| \leq \sum_{n=1}^{\infty} [\lambda_F(E_n) + \varepsilon 2^{-n}].$$

From this it follows that

$$\lambda_F(E) \leq \varepsilon + \sum_{n=1}^{\infty} [\lambda_F(E_n)]$$

and the subadditive property follows.

Exercise 753, page 340

One direction is easy. If β is full cover of (a, b) and we prune out intervals not inside (a, b) by writing $\beta' = \beta((a, b))$, then it is clear that

$$\lambda_F((a, b)) \leq \sup_{\pi \subset \beta'} \sum_{([u,v],w) \in \pi} |F(v) - F(u)| \leq V(F, [a, b]).$$

In the other direction ETC

Exercise 754, page 340

Suppose that $E \subset (-L, L)$ and that $|F'(x)| < M$ for all $x \in E$. Then

$$\beta = \left\{ ([u, v], w) : (u, v) \subset (-L, L), \frac{|F(v) - F(u)|}{v - u} < M \right\}$$

is a full cover of E . Thus, for any subpartition $\pi \subset \beta$,

$$\sum_{([u, v], w) \in \pi} |F(v) - F(u)| \leq 2ML.$$

It follows that $\lambda_F(E) \leq 2ML$ and so is finite.

Exercise 756, page 347

Develop the Henstock zero variation criterion for this integral and check that the usual zero derivative procedure will supply this fact.

Exercise 760, page 357

Note that there is no typo in the first inequality: the full variation is needed on the right-hand side. The second inequality, the easier to check, follows from the fact that the intersection of two full covers is again full. The first inequality follows from the fact that the intersection of two covering relations, one of which is full and the other fine, is again fine.

Exercise 765, page 357

If $f'(x) = 0$ for all $x \in E \setminus N$ show that $\lambda_f(E \setminus N) = 0$.

Exercise 767, page 357

It is enough to suppose that $\lambda_f(E) < \infty$. Let $C = \{x \in E : \lambda_f(\{x\}) > 0\}$; this set must include every point at which f fails to be continuous. Now let $C_n = \{x \in E : \lambda_f(\{x\}) > 1/n\}$. From measure-properties

$$\lambda_f(C_n) \leq \lambda_f(E) < \infty.$$

But if C_m contains k points then

$$\lambda_f(C_n) \geq k/n.$$

It follows that each set C_n is finite and hence the set C must be countable.

Exercise 768, page 358

Some authors would use the term weakly continuous at a point x_0 to mean that there is at least one sequence $c_n \rightarrow x_0$ and so that $|c_n - x_0| > 0$ and

$$f(c_n) - f(x_0) \rightarrow 0.$$

This condition is a little stronger than the definition in the text. For example the function $f(x) = 0$ if $x \neq 0$ and $f(0) = 1$ is weakly continuous at 0 in our sense but not in the stronger sense. The property in the exercise is dictated by the particular definition that we use for fine covers.

Here is a proof. Since f is weakly continuous at x_0 we know, by definition, that $\lambda_f^*(\{x_0\}) = 0$. For each integer n we can select a fine cover β_n of the set $\{x_0\}$ so that $V(\Delta f, \beta_n) < 1/n$. From β_n we can select a pair $([c_n, d_n], x_0)$ for which $d_n - c_n < 1/n$. Note that $c_n \leq x_0 \leq d_n$ and

$$|f(d_n) - f(c_n)| \leq V(\Delta f, \beta_n) < 1/n.$$

This pair of sequences $\{c_n\}$ and $\{d_n\}$ has all the properties that we need except they need not be monotonic. But there is a monotonic subsequence of the $\{c_n\}$ so that we can consider that we have selected that subsequence. Take a further subsequence so that both sequences are monotonic. The new sequences have all the properties that we need.

Exercise 769, page 358

Let

$$E = \{x : \liminf_{(I,x) \Rightarrow x} |\Delta f(I)| > 0\}$$

and

$$E = \{x : \liminf_{(I,x) \Rightarrow x} |\Delta f(I)| > 1/n\}.$$

The set of points where f is not weakly continuous is exactly the set $E = \bigcup_n E_n$. Note that $\beta = \{(I, x) : |\Delta f(I)| > 1/n\}$ is a full cover of E_n and apply the decomposition lemma from Section 5.1.8.

Exercise 771, page 358

Recall that F has finite variation on (a, b) if there is a number M and a full cover β of (a, b) so that

$$\sum_{([u,v],w) \in \pi} |F(v) - F(u)| \leq M$$

whenever π is a subpartition, $\pi \subset \beta$. If F has bounded variation on $[a, b]$ then certainly $M = V(F, [a, b])$ will work.

In the converse direction we suppose that M and β have been chosen with this property. For every subinterval $[c, d] \subset [a, b]$ there is a partition π contained in β for which evidently

$$|F(d) - F(c)| = \left| \sum_{([u,v],w) \in \pi} [F(v) - F(u)] \right| \leq \sum_{([u,v],w) \in \pi} |F(v) - F(u)| \leq M.$$

Fix some point x_0 in (a, b) and then we have the bound $|F(x)| \leq M + |F(x_0)|$ for every point x in (a, b) .

Now we estimate $V(F, [a, b])$. Take any choice of points

$$a = s_0 < s_1 < \cdots < s_{n-1} < s_n = b.$$

We note that

$$|F(s_1) - F(s_0)| \leq |F(a)| + M + |F(x_0)|$$

and that

$$|F(s_n) - F(s_{n-1})| \leq |F(b)| + M + |F(x_0)|.$$

We may choose a partition π from β so that π contains a partition of each of the remaining intervals $[s_1, s_2]$, $[s_2, s_3]$, $\dots$, $[s_{n-2}, s_{n-1}]$. This provides the inequality

$$\begin{aligned} \sum_{i=1}^n |F(s_i) - F(s_{i-1})| &\leq |F(a)| + M + |F(x_0)| + |F(b)| + M + |F(x_0)| + M = \\ &|F(a)| + |F(b)| + 3M + 2|F(x_0)|. \end{aligned}$$

This offers us an upper bound for $V(F, [a, b])$ and we have proved that F has bounded variation on $[a, b]$.

Exercise 773, page 358

If f is locally recurrent at every point of a set E then

$$\beta = \{(I, x) : \Delta f(I) = 0\}$$

is a fine cover of E . Thus

$$\lambda_f^*(E) \leq V(\Delta f, \beta) = 0.$$

Exercise 774, page 358

Define

$$\beta = \{(I, x) : \Delta f(I) \geq 0\}$$

and notice that this is a full cover of E . Apply the decomposition from Section 5.1.8 for β . There is an increasing sequence of sets $\{E_n\}$ with $E = \bigcup_{n=1}^{\infty} E_n$ and a sequence of compact intervals $\{I_{kn}\}$ covering E so that if x is any point in E_n and I is any subinterval of I_{kn} that contains x then (I, x) belongs to β .

We check that f is nondecreasing on each set $D_{nk} = E_n \cap I_{kn}$ in a certain strong way. For if either x or y belongs to the set D_{nk} and $[x, y] \subset I_{kn}$ then one of the pairs $([x, y], x)$ or $([x, y], y)$ belongs to β which requires that $f(x) \leq f(y)$.

Let $c = \inf D_{nk}$ and $d = \sup D_{nk}$. Suppose that $c = d$. Then D_{nk} contains a single point c and $\lambda_f(\{c\}) < \infty$, i.e., $\lambda_f(D_{nk}) < \infty$. Suppose instead that $c < d$. Let $D'_{nk} = D_{nk} \cap (c, d)$ so that D_{nk} contains, at most, two points c and d more than the set D'_{nk} . Let $\beta' = \beta[D_{nk}] \cap \beta((c, d))$. Then β' is a full cover of D'_{nk} . Let $\pi = \{\{[c_i, d_i], x_i\}\}$ be any subpartition contained in β' . We see from the manner in which f increases relative to the set D_{nk} that

$$\sum_i |f(d_i) - f(c_i)| \leq 2[f(d) - f(c)].$$

It follows that

$$\lambda_f(D'_{nk}) \leq V(\Delta f, \beta') \leq 2[f(d) - f(c)] < \infty.$$

Consequently,

$$\lambda_f(D_{nk}) \leq \lambda_f(D'_{nk}) + \lambda_f(\{c\}) + \lambda_f(\{d\}) < \infty$$

too, so that in either case $\lambda_f(D_{nk})$ is finite. It follows that λ_f is σ -finite on the set E since that set has been expressed as a union of a sequence of sets on each of which λ_f is σ -finite.

Exercise 775, page 359

Define three sets E_1 , E_2 , and E_3 . E_1 is the set of points at which f is locally nondecreasing. E_2 is the set of points at which $-f$ is locally nondecreasing. E_3 is the set of points at which f is locally recurrent. Since f is continuous it has the Darboux property. From that we see that $E_1 \cup E_2 \cup E_3 = \mathbb{R}$ since there are no other possibilities.

But $\lambda_f^*(E_3) = 0$ and λ_f is σ -finite on E_1 and E_2 (Exercise 774). It follows that the smaller measure λ_f^* must be σ -finite.

Exercise 776, page 359

Since the hint suggests that we can use Theorem 9.20 let us do so. There must be a sequence of compact sets $\{E_n\}$ covering E and a sequence of continuous functions of bounded variation $\{g_n\}$ so that f is Kolmogorov equivalent to g_n on E_n . In particular, we know that $g'_n(x)$ exists at almost every point. Therefore the set of points in E_n at which $f'(x) = g'_n(x)$ fails is a set of measure zero, say N_n . It follows that f is differentiable at every point of E with the possible exception of points in the measure zero set $\bigcup_{n=1}^{\infty} N_n$.

Exercise 778, page 365

Let

$$E = \{x : D^- f(x) < D_+ f(x)\}$$

and, for each rational number r , let

$$E_r = \{x : D^- f(x) < r < D_+ f(x)\}.$$

Note that E is the union of the countable collection of sets E_r taken over all rationals r . For each x in E_r there is a $\delta(x) > 0$ so that, for all $0 < h < \delta(x)$,

$$\Delta f([x-h, x]) < r\lambda([x-h, x])$$

and

$$\Delta f([x, x+h]) > r\lambda([x, x+h])$$

because of the values of the Dini derivatives.

Let

$$E_{rn} = \{x \in E : \delta(x) > 1/n\}$$

and check that

$$E_r = \bigcup_{n=1}^{\infty} E_{rn}.$$

We claim that, for each n , the set E_{rn} is countable. Indeed there cannot be two points x and y with $x < y$ in E_{rn} closer together than $1/n$. For if so, let $h = y - x$, note that $0 < h < \delta(x) < 1/n$ and $0 < h < \delta(y) < 1/n$. That would mean that

$$\Delta f([x, y]) < r\lambda([x, y]) < \Delta f([x, y])$$

which is impossible. Accordingly each E_{rn} is countable and so too also is E . The other set of the theorem can be handled by an identical proof.

Exercise 781, page 365

Consider first the set

$$A = \{x : D^- f(x) < D^+ f(x)\}$$

and, for each rational number r , let

$$A_r = \{x : D^- f(x) < r < D^+ f(x)\}.$$

Note that A is the union of the countable collection of sets A_r taken over all rational numbers r .

For each x in A_r we have $D^- f(x) < r$. Thus there is a $\delta(x) > 0$ so that, for all $0 < h < \delta(x)$,

$$f(x) - f(x-h) < rh.$$

For each $n = 1, 2, 3, \dots$ and each $k = 0, \pm 1, \pm 2, \dots$ write

$$A_{rnk} = \left[\frac{k-1}{n}, \frac{k}{n} \right] \cap \left\{ x \in A_r : \delta(x) > \frac{1}{n} \right\}.$$

Notice that

$$f(x) - f(y) < r(x-y)$$

for all $x < y$ with $x, y \in A_{rnk}$ and check that

$$A_r = \bigcup_{k=-\infty}^{\infty} \bigcup_{n=1}^{\infty} A_{rnk}.$$

Finally let E_{rnk} denote the closure of the set A_{rnk} . Each set E_{rnk} is compact and we claim that it contains no subinterval; in particular then it is a meager subset of $\mathbb{R}$.

Should such a set E_{rnk} contain an interval $[a, b]$ then, by the continuity of f we must conclude that the inequality stated above would require, for all $a < y < x < b$, that

$$f(x) - f(y) \leq r(x - y).$$

Consequently there would be no points y in (a, b) at which $r < D^+ f(y)$. But this is impossible since the set A_{rnk} is dense in the set E_{rnk} .

Thus we have displayed

$$A_r \subset \bigcup_{k=-\infty}^{\infty} \bigcup_{n=1}^{\infty} E_{rnk}$$

as a subset of a union of a sequence of meager subsets of $\mathbb{R}$.

It follows that the set A defined above is also a meager subset of $\mathbb{R}$. In a similar way we can conclude that each of the sets

$$\{x : D^- f(x) > D^+ f(x)\}$$

$$\{x : D_- f(x) > D_+ f(x)\}$$

and

$$\{x : D_- f(x) < D_+ f(x)\}$$

is a meager subset of $\mathbb{R}$. From this the theorem follows.

Exercise 783, page 366

Suppose that f is not nondecreasing on $[a, b]$. Then we can choose points $a \leq a' < b' \leq b$ with $f(b') < f(a')$. Thus $[f(b'), f(a')]$ is a nonempty compact subinterval of $[c, d]$. Take any y between $f(b')$ and $f(a')$. Let

$$M(y) = \sup\{x \in (a', b') : f(x) = y\}.$$

Check that $f(x) = y$ and that $D^+ f(x) \leq 0$ whenever $x = M(y)$. Thus, $y \in f(D)$. Consequently $f(D)$ contains $(f(b'), f(a'))$ and so, also, all compact subintervals contained in this open interval.

Exercise 784, page 366

We break the proof into a number of steps that follow Morse's original exposition.

Step 1. Suppose that f is strictly decreasing on a compact set $E \subset [a, b]$. If E contains no subinterval, then we claim that $f(E)$ is a compact subset of $[c, d]$ that also contains no interval.

We can define a strictly decreasing, continuous function $g : \mathbb{R} \rightarrow \mathbb{R}$ so that $f(x) = g(x)$ for all x in E by making g continuous and linear on all the open intervals complementary to E . We know that $f(E) = g(E)$ would be compact. Suppose, contrary to what we want, that $g(E)$ contains a subinterval J of $[c, d]$. We consider the inverse function g^{-1} which maps that subinterval J back into E . Such a function would be continuous and therefore maps J to some interval. That would require E to contain an interval.

Step 2. Define, for each integer $n = 1, 2, 3, \dots$,

$$E_n = \{x \in [a, b] : f(x+h) - f(x) \leq -h/n \text{ whenever } 0 \leq h \leq 1/n\}$$

Then we will prove that E_n is a compact subset of $[a, b]$ that contains no interval and that $f(E_n)$ is a compact subset of $[c, d]$ that contains no interval.

It is easy to check, using the continuity of f , that E_n is closed. Thus both E_n and $f(E_n)$ must be compact. We subdivide $[a, b]$ into a finite collection $\{J_k\}$ of compact, nonoverlapping subintervals of $[a, b]$, covering all of that interval and each of length less than $1/n$. It is easy to see that f is strictly decreasing on each set $J_k \cap E_n$. By our hypotheses the set A is dense in $[a, b]$ so that no one of these sets $J_k \cap E_n$ can contain an interval. In particular E_n itself can contain no interval. Moreover, by step 1, we conclude that $f(J_k \cap E_n)$ is a compact set that contains no subintervals of $[c, d]$. It follows that $f(E_n)$ is contained in the finite union of such sets and so must itself contain no subintervals of $[c, d]$.

Step 3. The set B is a meager subset of $[a, b]$ and the set $f(B)$ is a meager subset of $[c, d]$. This follows from step 2 since B is the union of the sequence of sets $\{E_n\}$ each of which is a meager subset of $[a, b]$, while $f(B)$ is the union of the sequence of sets $\{f(E_n)\}$, each of which is a meager subset of $[c, d]$.

Step 4. Suppose now that f is not nondecreasing on $[a, b]$. Then we can choose points $a \leq a' < b' \leq b$ with $f(b') < f(a')$. Thus $[f(b'), f(a')]$ is a nonempty compact subinterval of $[c, d]$. We know from the proof of the preliminary lemma that

f maps the set

$$D = \{x \in [a, b] : D^+ f(x) \leq 0\}$$

onto a set containing the open interval $(f(b'), f(a'))$. But we already have established that the set $f(B)$ is a meager subset of $[c, d]$. Using the fact that $B \cup C = D$, we conclude that $f(B) \cup f(C) = f(D) \supset (f(b'), f(a'))$. Thus $f(C)$ must contain a residual subset of the interval $[f(b'), f(a')]$.

Exercise 786, page 366

For example, consider the set

$$E = \{x : D^+ f(x) < r\}$$

and write, for positive integers m and n ,

$$E_{mn} = \{x : f(x+t) - f(x) - rt + t/m \leq 0 \text{ for all } 0 \leq t \leq 1/n\}.$$

Since f is continuous, we can check that each set E_{mn} is closed. But

$$E = \bigcup_{n=1}^{\infty} \bigcup_{m=1}^{\infty} E_{mn}$$

reveals that E must be Lebesgue measurable.

Exercise 789, page 367

Use the Darboux property of continuous functions. As a more challenging exercise the student may wish to prove this without the assumption of continuity.

Exercise 791, page 367

This follows immediately from Lemma 9.19 since we know (Exercise 775) that, for every continuous function f , the measure λ_f^* must be σ -finite.

Exercise 793, page 368

Let

$$\beta_1 = \{(I, x) : s\lambda(I) < |h(I, x)|\}$$

and

$$\beta_2 = \{(I, x) : |h(I, x)| < r\lambda(I)\}.$$

Note that β_1 is a fine cover of E and that β_2 is a full cover of E . Let β be any full cover of E and note that $\beta_1 \cap \beta$ is a fine cover of E and that $\beta_2 \cap \beta$ is a full cover of E . Thus

$$V^*(h, E) \leq V(h, \beta \cap \beta_2) \leq rV(\lambda, \beta \cap \beta_2) \leq rV(\lambda, \beta).$$

From this it follows that

$$V^*(h, E) \leq r\lambda^*(E).$$

Similarly

$$sV_*(\lambda, E) \leq V(s\lambda, \beta \cap \beta_1) \leq V(h, \beta \cap \beta_1) \leq V(h, \beta).$$

From this it follows that

$$s\lambda_*(E) \leq V^*(h, E).$$

Exercise 794, page 368

Use the methods already seen for Exercise 793.

Exercise 796, page 368

Write

$$E_n = \{x \in E : \overline{\text{lip}}_f(z) < n\}.$$

By Lemma 9.16

$$\lambda_f(E_n \cap [-n, n]) \leq n\lambda(E_n \cap [-n, n]) < \infty.$$

It follows that f has σ -finite variation in E . Note then, that if N is a null subset of E ,

$$\lambda_f(N) \leq \sum_{n=1}^{\infty} \lambda_f(E_n \cap N) \leq n\lambda(E_n \cap N) = 0.$$

This proves the final assertion.

Exercise 797, page 372

If S is a null set then $Z = S$ solves the exercise. Otherwise construct such a set by first taking a countable dense subset Z_1 of S . [The endpoints of the complementary intervals will suffice, unless S contains an interval. If S does contain an interval then include all rational numbers in that interval.] Now Z_1 is a countable subset of S and so has measure zero. For each integer n choose an open set G_n containing Z_1 with $\lambda(G_n) < 1/n$. Finally check that $Z = S \cap \bigcap_{n=1}^{\infty} G_n$ is a $\mathcal{G}$ -set and that $\lambda(Z) = \lambda(Z_1) = 0$.

Exercise 799, page 382

Take g' to denote the derivative of g where that exists and 0 otherwise; such a function is measurable and we will be able to apply Exercise 693.

Observe first that if $[c, d]$ is any compact interval then

$$|\Delta g([c, d])| \leq \int_{[c, d]} |g'(x)| dx.$$

This follows from the fact that g is continuous so that

$$\begin{aligned} |\Delta g([c, d])| &\leq \lambda(g([c, d])) \leq \lambda(g([c, d] \cap N)) + \lambda(g([c, d] \setminus N)) = \\ \lambda(g([c, d] \cap N)) &\leq \lambda_g([c, d] \cap N) = \int_{[c, d]} |g'(x)| dx. \end{aligned}$$

Use Theorem 9.7 and Theorem 9.23. Now apply Exercise 693 to obtain, for every $\varepsilon > 0$, a $\delta > 0$ so that if G is an open set with $\lambda(G) < \delta$ then

$$\int_G |g'(x)| dx < \varepsilon.$$

In particular, if we are given any sequence of nonoverlapping intervals $\{[c_n, d_n]\}$ for which $\sum_n \lambda([c_n, d_n]) < \delta$ then there

is an open set G covering these intervals for which $\lambda(G) < \delta$; it follows that

$$\sum_n |\Delta g([c_n, d_n])| \leq \sum_n \int_{[c_n, d_n]} |g'(x)| dx \leq \int_G |g'(x)| dx < \epsilon.$$

Exercise 800, page 387

Under these hypotheses there is an indefinite integral F of the function f on the open interval (a, b) . If F is uniformly continuous on (a, b) then we know that f is integrable on all of $[a, b]$. Thus it is enough to establish that when f has continuous upper and lower integrals on $[a, b]$ it follows that F is uniformly continuous on (a, b) .

Exercise 801, page 387

Define F appropriately, starting with

$$F(z) = \sum_{[a_i, b_i] \subset [a, z]} \int_{a_i}^{b_i} f(x) dx$$

for any $z \in E$ and, for $z \in (a_j, b_j)$, set

$$F(z) = F(a_j) + \int_{a_j}^z f(x) dx.$$

Obtain $V^*(\Delta F - f\lambda, [a, b]) = 0$ from $V^*(f\lambda, E) = 0$, $V^*(\Delta F, E) = 0$, and

$$V^*(\Delta F - f\lambda, [a, b] \setminus E) \leq \sum_{i=1}^{\infty} V^*(\Delta F - f\lambda, [a_i, b_i]).$$

Exercise 808, page 401

This is “intuitively obvious.” Certainly in dimension one, length is additive, in dimension two area is additive, in dimension three volume is additive, etc.

Well no. While the truth of the statement is hardly surprising and it is indeed trivial in dimension one, a proof would still be needed. Not all textbooks might supply such a proof but if you search enough there should be a number of

examples. McShane proves this as Lemma 2-1 (p. 255) of his text and includes the following comment:

In higher-dimensional spaces the result is still true, but the proof of that fact is tedious. Some people may think that this additivity is “intuitively evident” and that it is a waste of time to prove it. But even in the plane there are far more complicated dissections of an interval into subintervals than simple checkerboard patterns. . . . Besides that, who can honestly say that he has any clear-cut “intuitions” about 19-dimensional space?

E.J. McShane, *Unified Integration*, Academic Press (1983).

Exercise 809, page 401

This is the higher dimensional version of the Cousin lemma that was used extensively in the elementary chapters. As is usual in mathematics the higher dimension version can be proved by a similar method provided one takes the time to modify the argument as needed. The key tool in dimension one was the nested interval property asserting that a shrinking sequence of closed bounded intervals converged to a point. The same is true in higher dimensions. Having established this fact the proof of the Cousin lemma is then straightforward.

If you need to see a formal proof see Theorem 3-1, p. 258 in E. J. McShane, *Unified Integration*, Academic Press (1983). Henstock also takes the trouble to prove this assertion in detail in Theorem 4.1, p. 43 of R. Henstock, *Lectures on the Theory of Integration*, World Scientific (1988).

Exercise 818, page 408

Do not use Theorem 10.8 since that is not the point of the exercises.

Exercise 819, page 409

Do not use Theorem ?? since that is not the point of the exercise. You can use Lemma 10.2.

Exercise 823, page 413

This will require an application of the dominated convergence theorem. For details that can be used to prove this exercise as well as the preceding two exercises see the proof of Theorem 4–1, pp. 262–264 in E. J. McShane, *Unified Integration*, Academic Press (1983).

Exercise 825, page 414

This is given as Corollary 4–2, pp. 265–266 in E. J. McShane, *Unified Integration*, Academic Press (1983).

Exercise 826, page 414

You should be able to verify that

$$\int_{-1}^1 \left(\int_{-1}^1 f(x,y) dx \right) dy = \int_{-1}^1 \left(\int_{-1}^1 f(x,y) dy \right) dx = 0$$

while the double integral

$$\int \int_{[-1,1] \times [-1,1]} f(x,y) dx dy$$

fails to exist. For the double integral consider first the integrals over the squares $[n^{-1}, 1] \times [n^{-1}, 1]$ and $[-1, -n^{-1}] \times [-1, -n^{-1}]$ for $n = 2, 3, 4, \dots$

Exercise 827, page 415

You should be able to verify that

$$\int_0^1 \left(\int_{-1}^1 f(x,y) dy \right) dx = 0$$

but that both

$$\int_{-1}^1 \left(\int_0^1 f(x,y) dx \right) dy$$

and the double integral

$$\int \int_{[-1,1] \times [-1,1]} f(x,y) dx dy$$

fails to exist. For the double integral consider first the integrals over the intervals $[n^{-1}, 1] \times [0, 1]$ for $n = 2, 3, 4, \dots$

Exercise 828, page 415

You should be able to verify that

$$\int_0^1 \left(\int_{-1}^1 f(x,y) dy \right) dx = 1$$

but that

$$\int_{-1}^1 \left(\int_0^1 f(x,y) dx \right) dy = -1$$

while the double integral

$$\int \int_{[0,1] \times [0,1]} f(x,y) dx dy$$

fail to exist. For the double integral consider that the function $f(x,y) > 4/(27x^2)$ at every point in the set

$$\{(x,y) : n^{-1} \leq x \leq 1, 0 < y < x/2\}.$$

Exercise 829, page 415

Well you can indeed define anything you want but it needs to be consistent and useful. There are (see Exercise 827) situations in which only one of the expressions

$$\int_c^d \left(\int_a^b f(x,y) dx \right) dy$$

and

$$\int_a^b \left(\int_c^d f(x,y) dy \right) dx$$

exists. So which order do we take as our definition? There are also situations in which both exist (see Exercise 828) but have different values! This student's version of an integral would not even, then, allow us to rotate the axes by a right-angle without changing the integral radically. There are also situations in which both integrals exist and have the same value but the double integral does not exist in our sense (see Exercise 826) and shouldn't exist since it leads to unpleasant conclusions.

Exercise 830, page 418

Take any subdivision of $[a, b]$,

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

Let

$$E_i = \{x \in E : x_{i-1} < f(x) \leq x_i\}$$

and recall that

$$\mathcal{L}^n(E_i) = w(x_i) - w(x_{i-1}).$$

Check that

$$x_{i-1}[w(x_i) - w(x_{i-1})] \leq \int_{E_i} f(x) dx \leq x_i[w(x_i) - w(x_{i-1})].$$

This connects the Riemann sums

$$\sum_{i=1}^n x_i[w(x_i) - w(x_{i-1})] \quad \text{and} \quad \sum_{i=1}^n x_i[w(x_i) - w(x_{i-1})]$$

with the integral

$$\int_{\{x \in E : a < f(x) \leq b\}} f(x) dx = \sum_{i=1}^n \int_{E_i} f(x) dx.$$

Exercise 831, page 418

Remember that, the infinite integral

$$\int_{-\infty}^{\infty} s dw(s)$$

would be the same as

$$\lim_{n \rightarrow \infty} \int_{-n}^n s dw(s).$$

Thus we need to show that

$$\int_E f(x) dx = \lim_{n \rightarrow \infty} \int_{\{x \in E: -n < f(x) \leq n\}} f(x) dx$$

which is a simple measure-theoretic computation.

Exercise 832, page 422

Choose $z \in E$ but not in E_1 . Consider the intervals

$$I_n = (z - 1/n, z + 1/n).$$

If $E_1 \cap I_n \neq \emptyset$ for all n then from Exercise 23 we can deduce that z would have to belong to the closed set E_1 .

Exercise 833, page 422

If $n = 1$ this is obvious. Use induction on n .

Exercise 834, page 423

Suppose not, i.e., suppose that none of the sets E_k contain a portion of E . Then, using Exercise 833, select a portion $E \cap (a_1, b_1)$ so that $E \cap (a_1, b_1) = \emptyset$ and pass to a closed subinterval $[c_1, d_1] \subset (a_1, b_1)$ for which $E \cap [c_1, d_1] \neq \emptyset$. Continue inductively, choosing portions

$$E \cap (a_n, b_n) \subset [c_{n-1}, d_{n-1}]$$

and closed subintervals $[c_n, d_n] \subset (a_n, b_n)$ for which $E \cap [c_n, d_n] \neq \emptyset$.

The nested sequence of intervals $\{[c_n, d_n]\}$ all must contain a point z of E in common. But this point z cannot belong to any of the sets E_k which is in contradiction to the hypothesis that $E \subset \bigcup_{k=1}^{\infty} E_k$.

Exercise 835, page 423

To adjust the proof, at the n th stage of the induction select the interval $(a_n, b_n) \subset G_n$. The point z you will find must belong to each of the G_n and, consequently, to E . Sets of the form $E = \bigcap_{j=1}^{\infty} G_j$ for some sequence $\{G_j\}$ of open sets are said to be sets of type $\mathcal{G}_\delta$.

Index

- Abel test for uniform convergence, 104, 125, 600
- absolute continuity in Vitali's sense, 167, 181
- absolute continuity, 165, 220
- absolutely continuous, 356
- absolutely continuous function, 165, 220
- absolutely continuous function may not have bounded variation, 165, 220
- absolutely convergent integral, 63, 172, 234
- absolutely integrable, 90, 240
- absolutely integrable function, 95
- absolutely integrable function, 295
- accumulation point of view, 127
- algebraic number, 149
- almost everywhere, 206
- applications of the integral, 127
- approximation by Riemann sums, 173, 234
- area, 127
- arithmetic Construction of Cantor set, 157
- avoiding the mean-value theorem, 36
- backwards integral, 70
- Baire, René, 422
- Baire-Osgood Theorem, 422
- Banach, S., 210
- Banach, Stephan, 380
- Banach-Zarecki Theorem, 380
- Beppo Levi Theorem, 365
- Bers, Lipman, 511
- bijective function, 450
- Bliss theorem, 82
- Bliss, G. A., 82
- Bolzano-Weierstrass argument, 25, 26
- Bolzano-Weierstrass compactness argument, 20
- Borel family, 275
- bounded derived numbers, 40, 517
- bounded function, 23
- bounded integrable functions, 172
- bounded interval, 4
- bounded set, 4
- bounded variation, 91, 135
- boundedness properties of uniformly continuous functions, 22
- calculus integral, 60

- calculus integral is a nonabsolute integral, 91
- calculus student notation, 53, 72, 88, 121
- Cantor
 - set, 154, 157
 - theorem, 149
- Cantor dust, 154
- Cantor function, 158, 159
- Cantor function not absolutely continuous, 166, 222
- Cantor set, 157, 159
- Cantor set has measure zero, 202
- Cantor ternary set, 154
- Cantor, G., 149, 159
- Carathéodory characterization of measurable set, 286
- Carathéodory, Constantin, 288
- careless student, 41, 101, 120, 158, 205, 312
- Cartesian product, 450
- Cauchy
 - criterion for uniform convergence, 102
- Cauchy criteria, 242
- Cauchy criterion for series, 103
- Cauchy mean-value theorem, 35
- Cauchy-Schwarz inequality, 145, 610
- change of a function, 161
- change of variable, 54, 87, 176, 251
- characteristic function, 427
- characteristic function of the rationals, 16
- characterization of the integral, 388
- characterization of the Lebesgue integral, 293
- Ciesielski, K., 279
- class, *see* set
- classicalrealanalysis.com, 3
- closed interval, 4
- closed set, 4
- collection, *see* set
- complementary intervals of Cantor set, 156
- composition
 - of functions, 429
- connected set, 11
- constant functions, 38
- constant of integration, 43
- construction of Cantor function, 159
- construction of the integral, 285
- contiguous interval, 427
- continuity and zero variation, 164
- continuous upper and lower integrals, 386
- continuous upper/lower integrals, 386
- contraposition, 430
- convergence
 - uniform convergence, 102
- convergent integral, 63, 172, 234
- converges infinitely slowly, 576
- converse, 430
- converse to the mean-value theorem, 36
- countable number of discontinuities, 152
- countable set, 148, 149
- countable set has measure zero, 153
- counterexamples, 97
- Cousin covering argument, 20
- Cousin covering lemma, 194
- Cousin's lemma, 401

INDEX

- covering relation, 189
- curve, 134
- curve length, 344
- Darboux property, 25
- Darboux property of continuous functions, 26
- Darboux property of derivative, 37
- Darboux property of Dini derivatives, 366
- Darboux, J. G., 26, 37
- Darboux, Jean-Gaston, 238
- De Morgan's laws, 433
- decomposition, 195
- definite integral, 60, 169, 233
- definite vs. indefinite integrals, 72
- degenerate interval, 464
- Denjoy integral, 310
- Denjoy's program, 309
- Denjoy, Arnaud, 309, 310
- Denjoy, Arnoud, 350
- Denjoy-Perron integral, 84
- derivative, 28
 - Darboux property of, 37
 - not the limit of derivatives, 97
- derivative of the definite integral, 89, 177
- derivative of the indefinite integral, 55
- derivative of the integral, 253
- derivative of the limit is not the limit of the derivative, 97
- derivative of the variation, 94
- derivatives and uniform convergence, 111
- determined, 240
- devil's staircase, 158
- difference of sets, 460
- differentiable function is absolutely continuous, 166, 222
- differentiable functions are integrable, 66
- differentiable implies continuous, 29
- differentiation rules, 30
- Dini derivative, 360
- Dirichelet integral, 75, 548
- discontinuities of monotonic functions, 148
- discontinuity
 - of a limit of continuous functions, 97
- discontinuous limit of continuous functions, 97
- distance
 - between a point and a set, 16
- domain of a function, 435
- dummy variable, 71
- dx , 71
- empty set $\emptyset$, 435
- enumeration, 148
- epsilon, delta version of derivative, 28
- epsilons, 445
- equivalence relation, 435
- error estimate for Riemann sums, 79
- error estimate for Simpson's rule, 139
- error estimate for trapezoidal rule, 139
- exact computation by Riemann sums, 77
- existence of indefinite integrals, 48
- family, *see* set
- Fatou's lemma, 289

- fine cover, 190
- fine null, 209
- fine variational measure, 299, 351
- finite derived numbers, 166, 223
- finite variation, 356
- first mean-value theorem for integrals, 73
- fixed point, 26
- Flett's theorem, 37
- Flett, T., 512
- formula for the length of a curve, 135
- Freiling's criterion, 297
- Freiling, Chris, 49
- Freiling, Christopher, 297
- Fubini differentiation theorem, 228
- Fubini theorem, 416
- full cover, 190
- full null set, 207
- full variational measure, 299, 351
- function, 11
 - σ -finite variation, 356
 - absolutely continuous, 356
 - bijective, 450
 - bounded variation, 91
 - Cantor function, 159
 - characteristic function of a set, 427
 - composition, 429
 - distance between set and point, 16
 - domain of, 435
 - finite variation, 356
 - fixed point of, 26
 - injective, 450
 - inverse of, 440
 - Kolmogorov equivalence, 353, 356
 - Lipschitz, 40
 - Lipschitz numbers, 363
 - mutually singular, 356
 - one-to-one, 450
 - onto, 450
 - range of, 454
 - saltus, 356
 - singular, 356
 - smooth, 35
 - step function, 15
 - surjective, 450
 - Vitali property, 356
 - zero variation, 356
- function of Cantor, 158
- function with zero variation, 161
- growth lemmas, 364
- growth of a function, 161
- Harnack extension property, 387
- Harnack property, 387
- Heaviside function, 66
- Heaviside's function, 15
- Heine-Borel argument, 21
- Heine-Borel property, 26
- Hellinger integral, 347
- Hellinger, Ernst, 347
- Henstock property, 84

- Henstock's criterion, 173, 234
- Henstock, Ralph, 84, 173, 234, 350
- Henstock-Kurzweil integral, 180
- Heuer, Gerald A. , 569

- inadequate theory of integration, 184
- indefinite integral, 43
- indefinite integrals and bounded variation, 95
- indicator function, *see* characteristic function of a set
- indirect proof, 437
- inequalities, 175
- inequality
 - Cauchy-Schwarz, 610
- inequality properties of the integral, 85
- infinitely slowly, 576
- injective function, 450
- integrability on subintervals, 245
- integrable, 240
- integrable [calculus sense], 60
- integrable function, 169, 233, 240
- integral, 238
 - not the limit of the integral, 98
- integral inequalities, 64
- integral of nonnegative measurable function, 288
- integral of simple function, 288
- integral of the limit is not the limit of the integrals, 98
- integration by parts, 52, 87, 176, 253
- integration by substitution, 54, 87, 176
- integration of series, 256
- interchange of limit operations, 98
- intermediate value property, 26
 - of derivative, 37
- intersection of sets , 460
- intersection of two closed intervals, 4
- intersection of two open intervals, 4
- intervals, 3
- inverse
 - of a function, 440
- irrational number in the Cantor ternary set, 157
- Israel Halperin, 512
- iterated integral, 411

- Java Applets, 608
- Jordan decomposition, 93

- Kolmogorov equivalence, 353, 356
- Kolmogorov, A. N., 353
- Kurweil, Jaroslav, 84
- Kurzweil-Henstock integral, 180

- L'Hôpital's rule, 35
- law of the mean, 32
- least upper bound argument, 24
- least upper bound property, 26
- Lebesgue characterization of measurable, 286
- Lebesgue decomposition theorem, 293
- Lebesgue integrable, 181
- Lebesgue integral, 181
- Lebesgue measure, 264
- length of a curve, 134
- length of a graph, 134

- length of curve, 344
- Levi, Beppo, 365
- liminf comparison lemma, 368
- limit
 - interchange of limit operations, 98
- limitations of the calculus integral, 65, 151
- limits of Discontinuous Derivatives, 113
- limsup comparison lemma, 368
- linear combination, 14, 52
- linear combinations, 86, 175
- Lipschitz
 - condition, 40
 - function, 40
- Lipschitz condition, 24
- Lipschitz function, 50, 93, 166, 172, 222
- Lipschitz function is absolutely continuous, 166, 222
- Lipschitz numbers, 363
- local conditions for integrability, 382
- local integrability, 382
- locally bounded function, 23
- locally of bounded variation, 94
- locally strictly increasing function, 29
- logarithm, 145
- logarithm function, 100
- lower integral, 238
- lower Stieltjes integral, 315
- Lusin's conditions, 380
- Lusin, N., 350, 380
- Luzin, N. N., 575
- M-test, 104
- manipulations with indefinite integrals, 51
- Maple, 57, 143, 144
- mapping definition of continuity, 16
- maximum, 24
- McShane's characterization of the Lebesgue integral, 294
- McShane's criterion, 295
- McShane, E. J., 668
- McShane, Edward J., 295
- meager subset, 423
- mean-value theorem, 30, 32
 - second-order, 34, 35
- mean-value theorem for integrals, 73
- mean-value theorem of Cauchy, 35
- measurable function, 280, 407
- measurable set, 275, 406
- measure zero, 152, 200
- measure zero set, 202
- method of exhaustion, 128
- monotone convergence theorem, 114, 256
- Morse, Anthony P., 366
- mutually singular, 356
- mutually singular functions, 341
- nested interval property, 449
- no interval has measure zero, 153
- nonabsolutely integrable, 240
- nonabsolutely integrable function, 298
- nondifferentiable, 28
- nonmeasurable sets, 278

INDEX

- notation for indefinite integral, 47
- null function, 179
- numerical methods, 136
- one-to-one function, 450
- onto function, 450
- open interval, 4
- open set, 4
- ordered
 - pairs, 450
- oscillation, 241
- oscillation of a function, 17
- Osgood, W., 422
- Osgood-Baire Theorem, 422
- parametric curve, 134
- parametric equations, 134
- partial fractions, 57
- partition, 75, 188
- Pfeffer property, 370
- piecewise monotone, 37
- pointwise approximation by Riemann sums, 84, 173, 234
- pointwise continuous, 13
- portions, 422
- products of integrable functions, 69
- proof
 - contraposition, 430
 - converse, 430
 - indirect, 437
- properties of the definite integral, 85
- properties of the integral, 174
- properties of the total variation, 92
- pruning of a covering relation, 189
- quantifier
 - $\exists$, 453
 - $\forall$, 453
- range of a function, 454
- real numbers, 455
- rectifiable, 135
- reduction theorem, 327, 348
- relation, 455
 - equivalence relation, 435
- residual subset, 423
- review, 1
- Riemann criterion, 311
- Riemann integrable function, 183
- Riemann integral, 183
- Riemann sums, 75, 138, 173, 196, 234
- Riesz, Frigyes, 181
- Rolle's Theorem, 31
- Rozema, Edward , 607
- Saks, Stanislaw, 350
- saltus, 356
- second mean-value theorem for integrals, 74
- second mean-value theorem for the calculus integral, 145
- second order mean-value theorem, 34
- sequence definition of continuity, 16
- sequence of functions uniformly bounded, 107
- sequence of functions uniformly Cauchy, 102

sequence of functions uniformly convergent, 102

sequences and series of integrals, 96

sequences of functions of bounded variation, 94

sequences of sets of measure zero, 153

set

Cantor set, 154, 159

Cantor ternary set, 154, 157

Cartesian product, 450

countable, 148, 149

De Morgan's laws, 433

definition of, 458

difference of sets, 460

empty set $\emptyset$, 435

intersection of sets, 460

measure zero, 202

of ordered pairs, 450

of real numbers, 455

of zero content, 205

quantifier, 453

quantifier $\exists$, 453

quantifier $\forall$, 453

relation, 455

set-builder notation, 458

subset, 460

union of sets, 460

set of measure zero, 200

set-builder notation, 458

sets of measure zero, 152

simple function, 282

Simpson's rule, 139

sin integral function, 75, 548

singular, 356

singular function, 343

smooth function, 35

Solovay, R. M., 279

step function, 15

step functions, 49

step functions are integrable, 66

Stieltjes integral, 315, 417

Stieltjes, T. J., 315

straddled version of derivative, 29

subintervals, 87, 176

subpartition, 75, 153, 162, 189

subset, 460

summing inside the integral, 114, 256

surjective function, 450

Szökefalvi-Nagy, Béla, 181

tables of integrals, 58

TBB, 3, 181

ternary representation of Cantor set, 157

theorem

mean-value theorem, 32

of Cantor, 149

theorem of G. A. Bliss, 82

theorem of G. C. Young, 365

theorem of Morse, 366

theorem of W. H. Young, 365

total variation function, 91

transcendental number, 149

INDEX

- trapezoidal rule, 139
- trigonometric functions, 12
- Trillia Group, 151
- unbounded interval, 4
- unbounded limit of bounded functions, 97
- uniform Approximation by Riemann sums, 79
- uniform convergence, 102
 - Abel's test, 104, 600
 - Cauchy criterion, 102
 - Weierstrass M -test, 104
- uniform convergence and derivatives, 111
- uniform convergence of continuous functions, 109
- uniformly approximating the variation, 94
- uniformly bounded family of functions, 107
- uniformly Cauchy, 102
- uniformly continuous, 13
- uniformly full cover, 190
- uniformly full null, 209
- union of compact intervals, 212
- union of sets, 460
- unstraddled version of derivative, 29
- upper function, 49
- upper integral, 238
- upper Stieltjes integral, 315
- vanishing derivatives, 38
- vanishing derivatives with a few exceptions, 38
- variation expressed as a Stieltjes integral, 322
- variational measures, 299, 351
- Vitali continuous, 304
- Vitali cover, 190
- Vitali covering theorem, 210
- Vitali definition of absolute continuity, 167, 181
- Vitali property, 356
- Vitali property and differentiability, 372
- Vitali property characterized, 376
- Vitali's criterion, 304
- Vitali, G., 167, 181, 210, 278
- Vitali, Guiseppe, 303
- Weierstrass
 - M -test, 104
- why the finite exceptional set?, 65
- Young, Grace Chisolm, 365
- Young, William Henry, 365
- Zakon, Elias, 44
- Zakon, Elias, 151
- Zarecki, M. A., 380
- Zermelo–Fraenkel set theory, 279
- zero content, 205
- zero derivatives, 217
- zero derivatives imply zero variation, 164
- zero measure, 202
- zero variation, 161, 215, 356
- zero variation criterion, 318
- zero variation lemma, 164