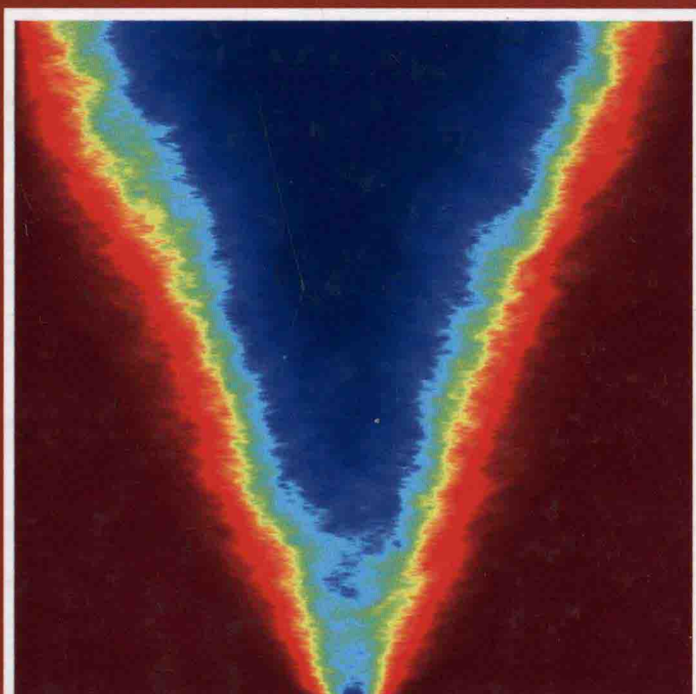
CAMBRIDGE TEXTS IN APPLIED MATHEMATICS

An Introduction to Computational Stochastic PDEs



GABRIEL J. LORD
CATHERINE E. POWELL
TONY SHARDLOW

AN INTRODUCTION TO COMPUTATIONAL STOCHASTIC PDES

GABRIEL J. LORD
Heriot-Watt University, Edinburgh

CATHERINE E. POWELL
University of Manchester

TONY SHARDLOW
University of Bath

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

32 Avenue of the Americas, New York NY 10013-2473, USA

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9780521899901

First published 2014

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data
Lord, Gabriel J., author.
An introduction to computational stochastic PDEs / Gabriel J. Lord, Heriot-Watt University,
Edinburgh, Catherine E. Powell, University of Manchester, Tony Shardlow, University of Bath.
pages cm – (Cambridge texts in applied mathematics; 50)
Includes bibliographical references and index.
ISBN 978-0-521-89990-1 (hardback) – ISBN 978-0-521-72852-2 (paperback)
1. Stochastic partial differential equations. I. Powell, Catherine E., author.
II. Shardlow, Tony, author. III. Title.
QA274.25.L67 2014
519.202–dc23 2014005535

ISBN 978-0-521-89990-1 Hardback
ISBN 978-0-521-72852-2 Paperback

Additional resources for this publication at www.cambridge.org/9780521899901

AN INTRODUCTION TO COMPUTATIONAL STOCHASTIC PDES

This book gives a comprehensive introduction to numerical methods and analysis of stochastic processes, random fields and stochastic differential equations, and offers graduate students and researchers powerful tools for understanding uncertainty quantification for risk analysis. Coverage includes traditional stochastic ordinary differential equations with white noise forcing, strong and weak approximation and the multilevel Monte Carlo method. Later chapters apply the theory of random fields to the numerical solution of elliptic PDEs with correlated random data, discuss the Monte Carlo method and introduce stochastic Galerkin finite element methods. Finally, stochastic parabolic PDEs are developed.

Assuming little previous exposure to probability and statistics, theory is developed in tandem with state-of-the-art computational methods through worked examples, exercises, theorems and proofs. The set of MATLAB codes included (and downloadable) allows readers to perform computations themselves and solve the test problems discussed. Practical examples are drawn from finance, mathematical biology, neuroscience, fluid flow modelling and materials science.

GABRIEL J. LORD is a professor in the Maxwell Institute, Department of Mathematics, at Heriot-Watt University. He has worked on stochastic PDEs and applications for more than ten years. He is the coeditor of *Stochastic Methods in Neuroscience* (with C. Laing) and has organised a number of international meetings in the field. He has served as an Associate Editor for the *SIAM Journal on Scientific Computing* and the *SIAM/ASA Journal on Uncertainty Quantification*.

CATHERINE E. POWELL is a senior lecturer in applied mathematics and numerical analysis at the University of Manchester. She has worked in the field of stochastic PDEs and uncertainty quantification for more than ten years. Together with Tony Shardlow, she initiated the NASPDE (Numerical Analysis of Stochastic PDEs) series of meetings. She has also served as an Associate Editor for the *SIAM/ASA Journal on Uncertainty Quantification*.

TONY SHARDLOW has been working in the numerical analysis group at the University of Bath since 2012. Before that, he held appointments at the universities of Manchester, Durham, Oxford and Minnesota. He completed his PhD in scientific computing and computational mathematics at Stanford University in 1997.

Cambridge Texts in Applied Mathematics

All titles listed below can be obtained from good booksellers or from Cambridge University Press. For a complete series listing, visit www.cambridge.org/mathematics.

Flow, Deformation and Fracture
G. I. BARENBLATT

The Mathematics of Signal Processing
STEVEN B. DAMELIN AND WILLARD MILLER, JR.

Nonlinear Dispersive Waves
MARK J. ABLOWITZ

Complex Variables: Introduction and Applications (2nd Edition)
MARK J. ABLOWITZ AND ATHANASSIOS S. FOKAS

Scaling
G. I. R. BARENBLATT

Introduction to Symmetry Analysis
BRIAN J. CANTWELL

Hydrodynamic Instabilities
FRANÇOIS CHARRU

A First Course in Continuum Mechanics
OSCAR GONZALEZ AND ANDREW M. STUART

Theory of Vortex Sound
M. S. HOWE

Applied Solid Mechanics
PETER HOWELL, GREGORY KOZYREFF AND JOHN OCKENDON

Practical Applied Mathematics: Modelling, Analysis, Approximation
SAM HOWISON

A First Course in the Numerical Analysis of Differential Equations (2nd Edition)
ARIEH ISERLES

A First Course in Combinatorial Optimization
JON LEE

An Introduction to Parallel and Vector Scientific Computation
RONALD W. SHONKWILER AND LEW LEFTON

Preface

Techniques for solving many of the differential equations traditionally used by applied mathematicians to model phenomena such as fluid flow, neural dynamics, electromagnetic scattering, tumour growth, telecommunications, phase transitions, etc. are now mature. Parameters within those models (e.g., material properties, boundary conditions, forcing terms, domain geometries) are often assumed to be known exactly, even when it is clear that is not the case. In the past, mathematicians were unable to incorporate noise and/or uncertainty into models because they were constrained both by the lack of computational resources and the lack of research into stochastic analysis. These are no longer good excuses. The rapid increase in computing power witnessed in recent decades allows the extra level of complexity induced by uncertainty to be incorporated into numerical simulations. Moreover, there are a growing number of researchers working on stochastic partial differential equations (PDEs) and their results are continually improving our theoretical understanding of the behaviour of stochastic systems. The transition from working with purely deterministic systems to working with stochastic systems is understandably daunting for recent graduates who have majored in applied mathematics. It is perhaps even more so for established researchers who have not received any training in probability theory and stochastic processes. We hope this book bridges this gap and will provide training for a new generation of researchers — that is, you.

This text provides a friendly introduction and practical route into the numerical solution and analysis of stochastic PDEs. It is suitable for mathematically grounded graduates who wish to learn about stochastic PDEs and numerical solution methods. The book will also serve established researchers who wish to incorporate uncertainty into their mathematical models and seek an introduction to the latest numerical techniques. We assume knowledge of undergraduate-level mathematics, including some basic analysis and linear algebra, but provide background material on probability theory and numerical methods for solving differential equations. Our treatment of model problems includes analysis, appropriate numerical methods and a discussion of practical implementation. MATLAB is a convenient computer environment for numerical scientific computing and is used throughout the book to solve examples that illustrate key concepts. We provide code to implement the algorithms on model problems, and sample code is available from the authors' or the publisher's website*. Each chapter concludes with exercises, to help the reader study and become more

* `http://www.cambridge.org/9780521728522`

familiar with the concepts involved, and a section of notes, which contains pointers and references to the latest research directions and results.

The book is divided into three parts, as follows.

Part One: Deterministic Differential Equations We start with a deterministic or non-random outlook and introduce preliminary background material on functional analysis, numerical analysis, and differential equations. Chapter 1 reviews linear analysis and introduces Banach and Hilbert spaces, as well as the Fourier transform and other key tools from Fourier analysis. Chapter 2 treats elliptic PDEs, starting with a two-point boundary-value problem (BVP), and develops Galerkin approximation and the finite element method. Chapter 3 develops numerical methods for initial-value problems for ordinary differential equations (ODEs) and a class of semilinear PDEs that includes reaction–diffusion equations. We develop finite difference methods and spectral and finite element Galerkin methods. Chapters 2 and 3 include not only error analysis for selected numerical methods but also MATLAB implementations for test problems that illustrate numerically the theoretical orders of convergence.

Part Two: Stochastic Processes and Random Fields Here we turn to probability theory and develop the theory of stochastic processes (one parameter families of random variables) and random fields (multi-parameter families of random variables). Stochastic processes and random fields are used to model the uncertain inputs to the differential equations studied in Part Three and are also the appropriate way to interpret the corresponding solutions. Chapter 4 starts with elementary probability theory, including random variables, limit theorems, and sampling methods. The Monte Carlo method is introduced and applied to a differential equation with random initial data. Chapters 5–7 then develop theory and computational methods for stochastic processes and random fields. Specific stochastic processes discussed include Brownian motion, white noise, the Brownian bridge, and fractional Brownian motion. In Chapters 6 and 7, we pay particular attention to the important special classes of stationary processes and isotropic random fields. Simulation methods are developed, including a quadrature scheme, the turning bands method, and the highly efficient FFT-based circulant embedding method. The theory of these numerical methods is developed alongside practical implementations in MATLAB.

Part Three: Stochastic Differential Equations There are many ways to incorporate stochastic effects into differential equations. In the last part of the book, we consider three classes of stochastic model problems, each of which can be viewed as an extension to a deterministic model introduced in Part One. These are:

Chapter 8	ODE (3.6)	+	white noise forcing
Chapter 9	Elliptic BVP (2.1)	+	correlated random data
Chapter 10	Semilinear PDE (3.39)	+	space–time noise forcing

Note the progression from models for time t and sample variable ω in Chapter 8, to models for space $\boldsymbol{x}$ and ω in Chapter 9, and finally to models for $t, \boldsymbol{x}, \omega$ in Chapter 10. In each case, we adapt the techniques from Chapters 2 and 3 to show that the problems are well posed and to develop numerical approximation schemes. MATLAB implementations are also discussed. Brownian motion is key to developing the time-dependent problems with white noise forcing considered in Chapters 8 and 10 using the Itô calculus. It is these types of differential equations that are traditionally known as stochastic differential equations (SODEs and SPDEs). In Chapter 9, however, we consider elliptic BVPs with both a forcing term and coefficients that are represented by random fields *not* of white noise type. Many authors prefer to reserve the term 'stochastic PDE' only for PDEs forced by white noise. We interpret it more broadly, however, and the title of this book is intended to incorporate PDEs with data and/or forcing terms described by both white noise (which is uncorrelated) and correlated random fields. The analytical tools required to solve these two types of problems are, of course, very different and we give an overview of the key results.

Chapter 8 introduces the class of stochastic ordinary differential equations (SODEs) consisting of ODEs with white noise forcing, discusses existence and uniqueness of solutions in the sense of Itô calculus, and develops the Euler–Maruyama and Milstein approximation schemes. Strong approximation (of samples of the solution) and weak approximation (of averages) are discussed, as well as the multilevel Monte Carlo method. Chapter 9 treats elliptic BVPs with correlated random data on two-dimensional spatial domains. These typically arise in the modelling of fluid flow in porous media. Solutions are also correlated random fields and, here, do not depend on time. To begin, we consider log-normal coefficients. After sampling the input data, we study weak solutions to the resulting deterministic problems and apply the Galerkin finite element method. The Monte Carlo method is then used to estimate the mean and variance. By approximating the data using Karhunen–Loève expansions, the stochastic PDE problem may also be converted to a deterministic one on a (possibly) high-dimensional parameter space. After setting up an appropriate weak formulation, the stochastic Galerkin finite element method (SGFEM), which couples finite elements in physical space with global polynomial approximation on a parameter space, is developed in detail. Chapter 10 develops stochastic parabolic PDEs, such as reaction–diffusion equations forced by a space–time Wiener process, and we discuss (strong) numerical approximation in space and in time. Model problems arising in the fields of neuroscience and fluid dynamics are included.

The number of questions that can be asked of stochastic PDEs is large. Broadly speaking, they fall into two categories: *forward problems* (sampling the solution, determining exit times, computing moments, etc.) and *inverse problems* (e.g., fitting a model to a set of observations). In this book, we focus on forward problems for specific model problems. We pay particular attention to elliptic PDEs with coefficients given by correlated random fields and reaction–diffusion equations with white noise forcing. We also focus on methods to compute individual samples of solutions and to compute moments (means, variances) of functionals of the solutions. Many other stochastic PDE models are neglected (hyperbolic problems, random domains, forcing by Levy processes, to name a few) as are many important questions (exit time problems, long-time simulation, filtering). However, this book covers a wide range of topics necessary for studying these problems and will leave the reader well prepared to tackle the latest research on the numerical solution of a wide range of stochastic PDEs.

Contents

1

Linear Analysis

This chapter introduces theoretical tools for studying stochastic differential equations in later chapters. §1.1 and §1.2 review Banach and Hilbert spaces, the mathematical structures given to sets of random variables and the natural home for solutions of differential equations. §1.3 reviews the theory of linear operators, especially the spectral theory of compact and symmetric operators, and §1.4 reviews Fourier analysis.

1.1 Banach spaces C^r and L^p

Banach and Hilbert spaces are fundamental to the analysis of differential equations and random processes. This section treats Banach spaces, reviewing first the notions of norm, convergence, and completeness before giving Definition 1.7 of a Banach space. We assume readers are familiar with real and complex vector spaces.

Definition 1.1 (norm) A *norm* $\|\cdot\|$ is a function from a real (respectively, complex) vector space X to $\mathbb{R}^+$ such that

(i) $\|u\| = 0$ if and only if $u = 0$,
(ii) $\|\lambda u\| = |\lambda|\,\|u\|$ for all $u \in X$ and $\lambda \in \mathbb{R}$ (resp., $\mathbb{C}$), and
(iii) $\|u + v\| \le \|u\| + \|v\|$ for all $u, v \in X$ (triangle inequality).

A *normed vector space* $(X, \|\cdot\|)$ is a vector space X with a norm $\|\cdot\|$. If only conditions (ii) and (iii) hold, $\|\cdot\|$ is called a *semi-norm* and denoted $|\cdot|_X$.

Example 1.2 $(\mathbb{R}^d, \|\cdot\|_2)$ is a normed vector space with

$$\|\boldsymbol{u}\|_2 := \left(|u_1|^2 + \cdots + |u_d|^2\right)^{1/2}$$

for the column vector $\boldsymbol{u} = [u_1, \dots, u_d]^{\mathsf{T}} \in \mathbb{R}^d$, where $|\cdot|$ denotes absolute value. More generally, $\|\boldsymbol{u}\|_\infty := \max\{|u_1|, \dots, |u_d|\}$ and $\|\boldsymbol{u}\|_p := (|u_1|^p + \cdots + |u_d|^p)^{1/p}$ for $p \ge 1$ is a norm and $(\mathbb{R}^d, \|\cdot\|_p)$ is a normed vector space. When $d = 1$, these norms are all equal to the absolute value.

Definition 1.3 (domain) A domain D is a non-trivial, connected, open subset of $\mathbb{R}^d$ and a domain is bounded if $D \subset \left\{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_2 \le R\right\}$ for some $R > 0$. The boundary of a domain is denoted ∂D and we always assume the boundary is piecewise smooth (e.g., the boundary of a polygon or a sphere).

Example 1.4 (continuous functions) For a subset $D \subset \mathbb{R}^d$, let $\mathrm{C}(D)$ denote the set of real-valued continuous functions on D. If D is a domain, functions in $\mathrm{C}(D)$ may be unbounded. However, functions in $\mathrm{C}(\bar{D})$, where $\bar{D}$ is the closure of D, are bounded and $(\mathrm{C}(\bar{D}), \|\cdot\|_\infty)$ is a normed vector space with the supremum norm,

$$\|u\|_\infty := \sup_{x \in \bar{D}} |u(x)|, \qquad u \in \mathrm{C}(\bar{D}).$$

A norm $\|\cdot\|$ on a vector space X measures the size of elements in X and provides a notion of convergence: for $u, u_n \in X$, we write $u = \lim_{n\to\infty} u_n$ or $u_n \to u$ as $n \to \infty$ in $(X, \|\cdot\|)$ if $\|u_n - u\| \to 0$ as $n \to \infty$. For example, the notion of convergence on $\mathrm{C}(\bar{D})$ is known as uniform convergence.

Definition 1.5 (uniform and pointwise convergence) We say $u_n \in \mathrm{C}(\bar{D})$ converges *uniformly* to a limit u if $\|u_n - u\|_\infty \to 0$ as $n \to \infty$. Explicitly, for every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that, for all $x \in \bar{D}$ and all $n \geq N$, $|u_n(x) - u(x)| < \epsilon$. In uniform convergence, N depends only on ϵ. This should be contrasted with the notion of pointwise convergence, which applies to all functions $u_n\colon D \to \mathbb{R}$. We say $u_n \to u$ *pointwise* if, for every $x \in D$ and every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|u_n(x) - u(x)| < \epsilon$. In pointwise convergence, N may depend both on ϵ and x.

There are many techniques, both computational and analytical, for finding approximate solutions $u_n \in X$ to mathematical problems posed on a vector space X. When u_n is a Cauchy sequence and X is complete, u_n converges to some $u \in X$, the so-called limit point, and this is often key in showing a mathematical model is well posed and proving the existence of a solution.

Definition 1.6 (Cauchy sequence, complete) Consider a normed vector space $(X, \|\cdot\|)$. A sequence $u_n \in X$ for $n \in \mathbb{N}$ is called a *Cauchy sequence* if, for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$\|u_n - u_m\| < \epsilon \qquad \text{for all } n, m \geq N.$$

A normed vector space $(X, \|\cdot\|)$ is said to be *complete* if every Cauchy sequence u_n in X converges to a limit point $u \in X$. In other words, there exists a $u \in X$ such that $\|u_n - u\| \to 0$ as $n \to \infty$.

Definition 1.7 (Banach space) A *Banach space* is a complete normed vector space.

Example 1.8 $(\mathbb{R}, |\cdot|)$ and $(\mathbb{R}^d, \|\cdot\|_p)$ for $1 \leq p \leq \infty$ are Banach spaces.

Example 1.9 $(\mathrm{C}(\bar{D}), \|\cdot\|_\infty)$ is a Banach space if D is bounded. See Exercise 1.1. If D is unbounded, the set of bounded continuous functions $\mathrm{C_b}(D)$ on D gives a Banach space.

The contraction mapping theorem is used in Chapters 3, 8, and 10 to prove the existence and uniqueness of solutions to initial-value problems.

Theorem 1.10 (contraction mapping) *Let Y be a non-empty closed subset of the Banach space $(X, \|\cdot\|)$. Consider a mapping $\mathcal{J}\colon Y \to Y$ such that, for some $\mu \in (0,1)$,*

$$\|\mathcal{J}u - \mathcal{J}v\| \leq \mu \|u - v\|, \qquad \text{for all } u, v \in Y. \tag{1.1}$$

There exists a unique fixed point of $\mathcal{J}$ in Y; that is, there is a unique $u \in Y$ such that $\mathcal{J}u = u$.

Proof Fix $u_0 \in Y$ and consider $u_n = \mathcal{J}^n u_0$ (the nth iterate of u_0 under application of $\mathcal{J}$). The sequence u_n is easily shown to be Cauchy in Y using (1.1) and therefore converges to a limit $u \in Y$ because Y is complete (as a closed subset of X). Now $u_n \to u$ and hence $u_{n+1} = \mathcal{J}u_n \to \mathcal{J}u$ as $n \to \infty$. We conclude that u_n converges to a fixed point of $\mathcal{J}$.

If $u, v \in Y$ are both fixed points of $\mathcal{J}$, then $\mathcal{J}u - \mathcal{J}v = u - v$. But (1.1) holds and hence $u = v$ and the fixed point is unique. □

Spaces of continuously differentiable functions

The smoothness, also called regularity, of a function is described by its derivatives and we now define spaces of functions with a given number of continuous derivatives. For a domain $D \subset \mathbb{R}^d$ and Banach space $(Y, \|\cdot\|_Y)$, consider a function $u\colon D \to Y$. We denote the partial derivative operator with respect to x_j by $\mathcal{D}_j := \frac{\partial}{\partial x_j}$. Given a multi-index $\alpha = (\alpha_1, \dots, \alpha_d)$, we define $|\alpha| := \alpha_1 + \cdots + \alpha_d$ and $\mathcal{D}^\alpha := \mathcal{D}_1^{\alpha_1} \cdots \mathcal{D}_d^{\alpha_d}$, so that

$$\mathcal{D}^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}.$$

Definition 1.11 (continuous functions)

(i) $\mathrm{C}(D,Y)$ is the set of continuous functions $u\colon D \to Y$. If D is bounded, we equip $\mathrm{C}(\bar{D},Y)$ with the norm

$$\|u\|_\infty := \sup_{x \in \bar{D}} \|u(x)\|_Y. \tag{1.2}$$

(ii) $\mathrm{C}^r(D,Y)$ with $r \in \mathbb{N}$ is the set of functions $u\colon D \to Y$ such that $\mathcal{D}^\alpha u \in \mathrm{C}(D,Y)$ for $|\alpha| \le r$; that is, functions whose derivatives up to and including order r are continuous. We equip $\mathrm{C}^r(\bar{D},Y)$ with the norm

$$\|u\|_{\mathrm{C}^r(\bar{D},Y)} := \sum_{0 \le |\alpha| \le r} \|\mathcal{D}^\alpha u\|_\infty.$$

We abbreviate the notation so that $\mathrm{C}(D,\mathbb{R})$ is denoted by $\mathrm{C}(D)$ and $\mathrm{C}^r(D,\mathbb{R})$ by $\mathrm{C}^r(D)$.

Proposition 1.12 *If the domain D is bounded,* $\mathrm{C}(\bar{D},Y)$ *and* $\mathrm{C}^r(\bar{D},Y)$ *are Banach spaces.*

Proof The case of $\mathrm{C}(\bar{D})$ is considered in Exercise 1.1. □

The following sets of continuous functions, which are not provided with a norm, are also useful.

Definition 1.13 (infinitely differentiable functions)

(i) $\mathrm{C}^\infty(D,Y)$ is the set $\bigcap_{r \in \mathbb{N}} \mathrm{C}^r(D,Y)$ of infinitely differentiable functions from D to Y.
(ii) $\mathrm{C}_\mathrm{c}^\infty(D,Y)$ is the set of $u \in \mathrm{C}^\infty(D,Y)$ such that $\operatorname{supp} u$ is a compact subset of D, where the support $\operatorname{supp} u$ denotes the closure of $\{x \in D\colon u(x) \ne 0\}$. (The definition of compact is recalled in Definition 1.66).

The spaces $\mathrm{C}^r(D,Y)$ specify the regularity of a function via the number, r, of continuous derivatives. More refined concepts of regularity include Hölder and Lipschitz regularity.

Definition 1.14 (Hölder and Lipschitz) Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces. A function $u \colon X \to Y$ is *Hölder continuous* with constant $\gamma \in (0,1]$ if there is a constant $L > 0$ so that

$$\|u(x_1) - u(x_2)\|_Y \le L\|x_1 - x_2\|_X^{\gamma}, \qquad \forall x_1, x_2 \in X.$$

If the above holds with $\gamma = 1$, then u is *Lipschitz continuous* or *globally* Lipschitz continuous to stress that L is uniform for $x_1, x_2 \in X$. The space $\mathrm{C}^{r,\gamma}(D)$ is the set of functions in $\mathrm{C}^r(D)$ whose rth derivatives are Hölder continuous with exponent γ.

Lebesgue integrals and measurability

The Lebesgue integral is an important generalisation of the Riemann integral. For a function $u \colon [a,b] \to \mathbb{R}$, the Riemann integral $\int_a^b u(x)\,dx$ is given by a limit of sums $\sum_{j=0}^{N-1} u(\xi_j)(x_{j+1} - x_j)$ for points $\xi_j \in [x_j, x_{j+1}]$, with respect to refinement of the partition $a = x_0 < \cdots < x_N = b$. In other words, u is approximated by a piecewise constant function, whose integral is easy to evaluate, and a limiting process defines the integral of u. The Lebesgue integral is also defined by a limit, but instead of piecewise constant approximations on a partition of $[a,b]$, approximations constant on *measurable sets* are used.

Let 1_F be the indicator function of a set F so

$$1_F(x) := \begin{cases} 1, & x \in F, \\ 0, & x \notin F. \end{cases}$$

Suppose that $\{F_j\}$ are measurable sets in $[a,b]$ (see Definition 1.15) and $\mu(F_j)$ denotes the measure of F_j (e.g., if $F_j = [a,b]$ then $\mu([a,b]) = |b - a|$). The Lebesgue integral of u with respect to the measure μ is defined via

$$\int_a^b u(x)\,d\mu(x) = \lim \sum_j u_j\, \mu(F_j),$$

where the limit is taken as the functions $\sum_j u_j 1_{F_j}(x)$ converge to $u(x)$. The idea is illustrated in Figure 1.1, where the function $u(x)$ is approximated by $\sum_{i=1}^{3} u_j 1_{F_j}(x)$ for

$$\begin{aligned} F_1 &= u^{-1}([-1,-0.5]), & F_2 &= u^{-1}((-0.5,0.5]), & F_3 &= u^{-1}((0.5,1]), \\ u_1 &= -0.8, & u_2 &= 0, & u_3 &= 0.8. \end{aligned}$$

Here, $u^{-1}([a,b]) := \{x \in \mathbb{R} : u(x) \in [a,b]\}$. To precisely define the Lebesgue integral, we must first form a collection of subsets $\mathcal{F}$ that we can measure.

Definition 1.15 (σ-algebra) A set $\mathcal{F}$ of subsets of a set X is a *σ-algebra* if

(i) the empty set $\{\} \in \mathcal{F}$,
(ii) the complement $F^{\mathrm{c}} := \{x \in X : x \notin F\} \in \mathcal{F}$ for all $F \in \mathcal{F}$, and
(iii) the union $\cup_{j \in \mathbb{N}} F_j \in \mathcal{F}$ for $F_j \in \mathcal{F}$.