Grundlehren der mathematischen Wissenschaften 352
A Series of Comprehensive Studies in Mathematics

Scott Armstrong
Tuomo Kuusi
Jean-Christophe Mourrat

Quantitative Stochastic Homogenization and Large-Scale Regularity

Springer

Grundlehren der mathematischen Wissenschaften

A Series of Comprehensive Studies in Mathematics

Volume 352

Grundlehren der mathematischen Wissenschaften (subtitled Comprehensive Studies in Mathematics), Springer's first series in higher mathematics, was founded by Richard Courant in 1920. It was conceived as a series of modern textbooks. A number of significant changes appear after World War II. Outwardly, the change was in language: whereas most of the first 100 volumes were published in German, the following volumes are almost all in English. A more important change concerns the contents of the books. The original objective of the Grundlehren had been to lead readers to the principal results and to recent research questions in a single relatively elementary and accessible book. Good examples are van der Waerden's 2-volume Introduction to Algebra or the two famous volumes of Courant and Hilbert on Methods of Mathematical Physics.

Today, it is seldom possible to start at the basics and, in one volume or even two, reach the frontiers of current research. Thus many later volumes are both more specialized and more advanced. Nevertheless, most books in the series are meant to be textbooks of a kind, with occasional reference works or pure research monographs. Each book should lead up to current research, without over-emphasizing the author's own interests. There should be proofs of the major statements enunciated, however, the presentation should remain expository. Examples of books that fit this description are Maclane's Homology, Siegel & Moser on Celestial Mechanics, Gilbarg & Trudinger on Elliptic PDE of Second Order, Dafermos's Hyperbolic Conservation Laws in Continuum Physics ... Longevity is an important criterion: a GL volume should continue to have an impact over many years.Topics should be of current mathematical relevance, and not too narrow.

The tastes of the editors play a pivotal role in the selection of topics.

Authors are encouraged to follow their individual style, but keep the interests of the reader in mind when presenting their subject. The inclusion of exercises and historical background is encouraged.

The GL series does not strive for systematic coverage of all of mathematics. There are both overlaps between books and gaps. However, a systematic effort is made to cover important areas of current interest in a GL volume when they become ripe for GL-type treatment.

As far as the development of mathematics permits, the direction of GL remains true to the original spirit of Courant. Many of the oldest volumes are popular to this day and some have not been superseded. One should perhaps never advertise a contemporary book as a classic but many recent volumes and many forthcoming volumes will surely earn this attribute through their use by generations of mathematicians.

More information about this series at http://www.springer.com/series/138

Scott Armstrong · Tuomo Kuusi ·
Jean-Christophe Mourrat

Quantitative Stochastic Homogenization and Large-Scale Regularity

Springer

Scott Armstrong
Courant Institute of Mathematical Sciences
New York, NY, USA

Jean-Christophe Mourrat
Département de Mathématiques
École Normale Supérieure
Paris, France

Tuomo Kuusi
The Department of Mathematics and
Statistics, Faculty of Science
University of Helsinki
Helsinki, Finland

ISSN 0072-7830 ISSN 2196-9701 (electronic)
Grundlehren der mathematischen Wissenschaften
ISBN 978-3-030-15544-5 ISBN 978-3-030-15545-2 (eBook)
https://doi.org/10.1007/978-3-030-15545-2

Library of Congress Control Number: 2019936291

Mathematics Subject Classification (2010): 35B27, 60F17, 35B65, 35R60, 35J25, 60H15

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Many microscopic models lead to partial differential equations with rapidly oscillating coefficients. A particular example, which is the main focus of this book, is the scalar, uniformly elliptic equation

$$-\nabla \cdot (\mathbf{a}(x)\nabla u) = f, \tag{0.1}$$

where the interest is in the behavior of the solutions on length scales much larger than the unit scale (the *microscopic* scale on which the coefficients are varying). The coefficients are assumed to be valued in the positive definite matrices, and may be periodic, almost periodic, or stationary random fields. Such equations arise in a variety of contexts such as heat conduction and electromagnetism in heterogeneous materials, or through their connection with stochastic processes.

To emphasize the highly heterogeneous nature of the problem, it is customary to introduce a parameter $0 < \varepsilon \ll 1$ to represent the ratio of the microscopic and macroscopic scales. The equation is then rescaled as

$$-\nabla \cdot \left(\mathbf{a}\left(\frac{x}{\varepsilon}\right)\nabla u^{\varepsilon}\right) = f, \tag{0.2}$$

with the problem reformulated as that of determining the asymptotic behavior of u^{ε}, subject to appropriate boundary conditions, as $\varepsilon \to 0$.

It has been known since the early 1980s that, under very general assumptions, the solution u^{ε} of the heterogeneous equation converges in L^2 to the solution u of a constant-coefficient equation

$$-\nabla \cdot (\bar{\mathbf{a}}\nabla u) = f. \tag{0.3}$$

We call this the *homogenized equation* and the coefficients the *homogenized* or *effective coefficients*. The matrix $\bar{\mathbf{a}}$ will depend on the coefficients $\mathbf{a}(\cdot)$ in a very complicated fashion: there is no simple formula for $\bar{\mathbf{a}}$ except in dimension $d = 1$ and some special situations in dimension $d = 2$. However, if one is willing to perform the computational work of approximating the homogenized coefficients and to tolerate the error in replacing u^ε by u, then there is a potentially huge payoff to be gained in terms of a reduction of the complexity of the problem. Indeed, up to a change of variables, the homogenized equation is simply the Poisson equation, which can be numerically computed in linear time and memory and is obviously independent of $\varepsilon > 0$. In contrast, the cost of computing the solution to the heterogeneous equation explodes as ε becomes small, and can be considered out of reach.

There is a vast and rich mathematical literature on homogenization developed in the last 40 years and already many good expositions on the topic (see, for instance, the books [5, 24, 30, 38, 39, 81, 87, 116, 123, 125]). Most of these works are focused on qualitative results, such as proving the existence of a homogenized equation which characterizes the limit as $\varepsilon \to 0$ of solutions. The need to develop efficient methods for determining $\bar{\mathbf{a}}$ and for estimating the error in the homogenization approximation (e.g., $\|u^\varepsilon - u\|_{L^2}$) motivates the development of a *quantitative* theory of homogenization. However, until recently, nearly all of the quantitative results were confined to the rather restrictive case of periodic coefficients. The main reason for this is that quantitative homogenization estimates in the periodic case are vastly simpler to prove than under essentially any other hypothesis (even the almost periodic case). Indeed, the problem can be essentially reduced to one on the torus and compactness arguments then yield optimal estimates. In other words, in the periodic setting, the typical arguments of qualitative homogenization theory can be made quantitative in a relatively straightforward way.

This book is concerned with the quantitative theory of homogenization for nonperiodic coefficient fields, focusing on the case in which $\mathbf{a}(x)$ is a stationary random field satisfying quantitative ergodicity assumptions. This is a topic which has undergone a rapid development since its birth at the beginning of this decade, with new results and more precise estimates coming at an ever accelerating pace. Very recently, there has been a convergence toward a common philosophy and set of core ideas, which has resulted in a complete and optimal theory. This book gives a complete and self-contained presentation of this theory.

We have written it with several purposes and audiences in mind. Experts on the topic will find new results as well as arguments which have been greatly simplified compared to the previous state of the literature. Researchers interested in stochastic homogenization will hopefully find a useful reference to the main results in the field and a roadmap to the literature. Our approach to certain topics, such as the construction of the Gaussian free field or the relation between Sobolev norms and the heat kernel, could be of independent interest to certain segments of the probability and analysis communities. We have written the book with newcomers to homogenization in mind and, most of all, graduate students and young researchers. In particular, we expect that readers with a basic knowledge of probability and analysis, but perhaps without expertise in elliptic regularity, the Gaussian free field, negative and fractional Sobolev spaces, etc, should not have difficulty following the flow of the book. These topics are introduced as they arise and are developed in a mostly self-contained way.

Before we give a summary of the topics we cover and the approach we take, let us briefly recall the historical and mathematical context. In the case of stationary random coefficients, there were very beautiful, soft arguments given independently in the early 1980s by Kozlov [88], Papanicolaou and Varadhan [114], and Yurinski [131] which give proofs of qualitative homogenization under very general hypotheses. A few years later, Dal Maso and Modica [40, 41] extended these results to nonlinear equations using variational arguments inspired by Γ-convergence. Each of the proofs in these papers relies in some way on an application of the ergodic theorem applied to the gradient (or energy density) of certain solutions of the heterogeneous equation. In order to obtain a convergence rate for the limit given by the ergodic theorem, it is necessary to verify quantitative ergodic conditions on the underlying random sequence or field. It is therefore necessary and natural to impose such a condition on the coefficient field $\mathbf{a}(x)$. However, even under the strongest of mixing assumptions (such as the finite range of dependence assumption we work with for most of this book), one faces the difficulty of transferring the quantitative ergodic information contained in these strong mixing properties from the coefficients to the solutions themselves, since the ergodic theorem must be applied to the latter. This is difficult because, of course, the solutions depend on the coefficient field in a very complicated, nonlinear, and nonlocal way.

Gloria and Otto [72, 73] were the first to address this difficulty in a satisfactory way in the case of coefficient fields that can be represented as functions of countably many independent random variables. They used an idea from statistical mechanics, previously introduced in the context of homogenization by Naddaf and Spencer [109], of viewing the solutions as functions of these independent random variables and applying certain general concentration inequalities such as the Efron–Stein or logarithmic Sobolev inequalities. If one can quantify the dependence of the solutions on a resampling of each independent random variable, then these inequalities immediately give bounds on the fluctuations of solutions. Gloria and Otto used this method to derive estimates on the first-order correctors which are optimal in terms of the ratio of length scales (although not optimal in terms of stochastic integrability).

The point of view developed in this book is different and originates in works of Armstrong and Smart [18], Armstrong and Mourrat [14], and the authors [11, 12]. Rather than study solutions of the equation directly, the main idea is to focus on certain energy quantities, which allow us to implement a progressive coarsening of the coefficient field and capture the behavior of solutions on large—but finite—length scales. The approach can thus be compared with renormalization group arguments in theoretical physics. The core of the argument is to establish that, as we pass to larger and larger length scales, these energy quantities become essentially local, additive functions of the coefficient field. It is then straightforward to optimally transfer the mixing properties of the coefficients to the energy quantities and then to the solutions. This perspective has motivated numerous subsequent developments by many researchers and we refer to the discussions at the end of each chapter for more precise references.

The quantitative analysis of the energy quantities is the focus of the first part of the book. After the first introductory chapter, the strategy naturally breaks into several distinct steps:

- Obtaining an *algebraic rate of homogenization* (Chap. 2): roughly speaking, we show that there is a positive exponent α such that the solutions u^ε and u of (0.2) and (0.3), respectively, are apart by $O(\varepsilon^\alpha)$. Here, the emphasis is on obtaining estimates with optimal stochastic integrability, while the exponent α representing the scaling of the error is clearly suboptimal. A precise statement can be found in Theorem 2.18, the proof of which is based on the subadditive and convex analytic structure endowed by the variational formulation of the equation. The basic idea is that, rather than attempting to understand the solutions directly, we should first analyze the behavior of their *energy densities*—or, to be precise, local, subadditive quantities related to their energy densities—which turn out to be better behaved. Specifically, we obtain a convergence rate for these subadditive quantities and show that this, in turn, implies a quantitative rate of homogenization for the Dirichlet problem.
- Establishing a *large-scale regularity theory* (Chap. 3): we show that solutions of an equation with stationary random coefficients are much more regular than one can show by naively applying deterministic elliptic regularity estimates. We prove this by showing that the extra regularity is inherited from the homogenized equation by approximation, using a Campanato-type iteration and the quantitative homogenization results obtained in the previous chapter.
- Deriving *optimal quantitative estimates* for the first-order correctors (Chap. 4): with the regularity theory in place, we can turn our attention to improving the scaling of the homogenization error. It is both natural and convenient (and for many purposes sufficient) to focus on the behavior of the first-order correctors ϕ_e. These are the functions for which $x \mapsto e \cdot x + \phi_e(x)$ is a global solution of (0.1) with $f = 0$ which stays close to the affine function $e \cdot x$. By the naive, classical two-scale expansion, we should expect that for a general solution u^ε of (0.2),

$$u^\varepsilon(x) \simeq w^\varepsilon(x) := u(x) + \varepsilon \sum_{k=1}^{d} \partial_{x_j} u(x) \phi_{e_k}\left(\frac{x}{\varepsilon}\right), \tag{0.4}$$

  see Fig. 1.2. Good quantitative information on the correctors therefore gives good quantitative information about homogenization more generally. By implementing a modification of the renormalization scheme of Chap. 2, with the major additional ingredient of the large-scale regularity theory, we gradually improve the exponent for convergence of the energy densities of the first-order correctors from a very tiny $\alpha > 0$ to $\alpha = \frac{d}{2}$, the optimal exponent predicted by the scaling of the central limit theorem. Roughly speaking, we show that, for every $\psi \in C_c^\infty(\mathbb{R}^d)$ with $\int \psi = 1$,

$$\left|\frac{1}{2}e\cdot\bar{\mathbf{a}}e-\int_{\mathbb{R}^d}\psi(x)\left(\frac{1}{2}\left(e+\nabla\phi_e\left(\frac{x}{\varepsilon}\right)\right)\cdot\mathbf{a}\left(\frac{x}{\varepsilon}\right)\left(e+\nabla\phi_e\left(\frac{x}{\varepsilon}\right)\right)\right)dx\right|=O(\varepsilon^{\frac{d}{2}}). \tag{0.5}$$

For a random field f^ε with range of dependence ε and unit variance, the standard deviation of the quantity $\int_{\mathbb{R}^d}\psi(x)f^\varepsilon(x)dx$ should scale the same way as a sum of $O(\varepsilon^{-d})$ many independent random variables of unit variance, which is $O(\varepsilon^{\frac{d}{2}})$ and matches (0.5). Once we have proved this estimate for the energy densities, we can read off a complete and optimal quantitative description of the behavior of the correctors themselves: see Theorems 4.1 and 4.24 for the precise statements. As in Chap. 2, in keeping with the spirit of the renormalization ideas, the quantities in (0.5) are studied only indirectly and the focus is rather on more local quantities which can be thought of as coarsenings of the coefficient field.

- We go beyond the optimal quantitative estimates to a description of the *next-order behavior* of the first-order correctors (Chap. 5). That is, we characterize the fluctuations of the energy densities of the first-order correctors by proving their convergence to white noise; consequently, we obtain the scaling limit of the first-order correctors to a generalized Gaussian free field.
- Combining the optimal estimates on the first-order correctors with classical arguments from homogenization theory, we obtain *optimal quantitative estimates on the homogenization error*, and the two-scale expansion, for Dirichlet and Neumann boundary value problems (Chap. 6). What we roughly show is that, for given sufficiently smooth data, the solutions u^ε and u of (0.2) and (0.3), respectively, and w^ε in (0.4) satisfy the estimates

$$\|u^\varepsilon-u\|_{L^2}=O(\varepsilon)\quad\text{and}\quad\|\nabla u^\varepsilon-\nabla w^\varepsilon\|_{L^2}=O(\varepsilon^{\frac{1}{2}}),$$

with an extra factor of $|\log\varepsilon|^{\frac{1}{2}}$ in dimension $d=2$ (which is intrinsic). See Theorems 6.11 and 6.17 for the precise statements of these estimates, which agree with the classical estimates in the case of periodic coefficients (up to the logarithmic correction in two dimensions).

These six chapters represent, in our view, the essential part of the theory. The first four chapters should be read consecutively (Sects. 3.5 and 3.6 can be skipped), while Chaps. 5 and 6 are independent of each other.

Chapter 7 complements the regularity theory of Chap. 3 by developing local and global gradient L^p estimates $(2<p<\infty)$ of Calderón–Zygmund-type for equations with right-hand side. Using these estimates, in Sect. 7.3 we extend the results of Chap. 6 by proving optimal quantitative bounds on the error of the two-scale expansion in $W^{1,p}$-type norms. Except for the last section, which requires the optimal bounds on the first-order correctors proved in Chap. 4, this chapter can be read after Chap. 3.

Chapter 8 extends the analysis to the time-dependent parabolic equation

$$\partial_t u - \nabla \cdot \mathbf{a}\nabla u = 0.$$

The main focus is on obtaining a suboptimal error estimate for the Cauchy–Dirichlet problem and a parabolic version of the large-scale regularity theory. Here the coefficients $\mathbf{a}(x)$ depend only on space, and the arguments in the chapter rely on the estimates on first-order correctors obtained in Chaps. 2 and 3 in addition to some relatively routine deterministic arguments. In Chap. 8, we also prove decay estimates on the elliptic and parabolic Green functions as well as on their derivatives, homogenization error and two-scale expansions.

In Chap. 9, we study the decay, as $t \to \infty$, of the solution $u(t,x)$ of the parabolic initial-value problem

$$\begin{cases} \partial_t u - \nabla \cdot (\mathbf{a}\nabla u) = 0 & \text{in } (0,\infty) \times \mathbb{R}^d, \\ u(0,\cdot) = \nabla \cdot \mathbf{g} & \text{on } \mathbb{R}^d, \end{cases}$$

where $\mathbf{g}$ is a bounded, stationary random field with a unit range of dependence. We show that the solution u decays to zero at the same rate as one has in the case $\mathbf{a} = \mathsf{Id}$. As an application, we upgrade the quantitative homogenization estimates for the parabolic and elliptic Green functions to the optimal scaling (see Theorem 9.11 and Corollary 9.12).

In Chap. 10, we show how the variational methods in this book can be adapted to non-self-adjoint operators, in other words, linear equations with nonsymmetric coefficients. In Chap. 11, we give a generalization to the case of nonlinear equations. In particular, in both of these chapters we give a full generalization of the results of Chaps. 1 and 2 to these settings, as well as the large-scale $C^{0,1}$ estimate of Chap. 3.

We would like to thank several of our colleagues and students for their helpful comments, suggestions, and corrections: Alexandre Bordas, Sanchit Chaturvedi, Paul Dario, Sam Ferguson, Chenlin Gu, Jan Kristensen, Jules Pertinand, Christophe Prange, Armin Schikorra, Charlie Smart, Tom Spencer, Stephan Wojtowytsch, Wei Wu, and Ofer Zeitouni. We particularly thank Antti Hannukainen for his help with the numerical computations that generated Fig. 5.3. SA was partially supported by NSF Grant DMS-1700329. TK was partially supported by the Academy of Finland and he thanks Giuseppe Mingione for the invitation to give a graduate course at the University of Parma. JCM was partially supported by the ANR grant LSD (ANR-15-CE40-0020-03).

We expect that, despite our best efforts, some slight inaccuracies and typos remain in the manuscript. We encourage readers to send us any they may find by email. We will maintain a list of typos and misprints found after publication on our webpages.

New York, USA — Scott Armstrong
Helsinki, Finland — Tuomo Kuusi
Paris, France — Jean-Christophe Mourrat
January 2019

Contents

Assumptions and Examples

We present the assumptions which are in force throughout (most of) the book and give concrete examples of coefficient fields satisfying them.

Assumptions

Except where specifically indicated otherwise, the following standing assumptions are in force throughout the book.

We fix a constant $\Lambda > 1$ called the *ellipticity constant*, and a dimension $d \geqslant 2$.

We let Ω denote the set of all measurable maps $\mathbf{a}(\cdot)$ from $\mathbb{R}^d$ into the set of symmetric $d \times d$ matrices, denoted by $\mathbb{R}^{d\times d}_{\mathrm{sym}}$, which satisfy the uniform ellipticity and boundedness condition

$$|\xi|^2 \leqslant \xi \cdot \mathbf{a}(x)\xi \leqslant \Lambda |\xi|^2, \quad \forall \xi \in \mathbb{R}^d. \tag{0.6}$$

That is,

$$\Omega := \left\{ \mathbf{a} : \mathbf{a} \text{ is a Lebesgue measurable map from } \mathbb{R}^d \textit{ to } \mathbb{R}^{d\times d}_{\mathrm{sym}} \text{ satisfying } (0.6) \right\}. \tag{0.7}$$

The entries of an element $\mathbf{a} \in \Omega$ are written as $\mathbf{a}_{ij}$, $i, j \in \{1, \ldots, d\}$.

We endow Ω with a family of σ-algebras $\{\mathcal{F}_U\}$ indexed by the family of Borel subsets $U \subseteq \mathbb{R}^d$, defined by

$$\begin{aligned} \mathcal{F}_U := &\text{ the } \sigma - \text{algebra generated by the following family :} \\ &\left\{ \mathbf{a} \mapsto \int_{\mathbb{R}^d} \mathbf{a}_{ij}(x)\varphi(x)dx : \varphi \in C_c^\infty(U), i, j \in \{1, \ldots, d\} \right\}. \end{aligned} \tag{0.8}$$

The largest of these σ-algebras is denoted $\mathcal{F} := \mathcal{F}_{\mathbb{R}^d}$. For each $y \in \mathbb{R}^d$, we let $T_y : \Omega \to \Omega$ be the action of translation by y,

$$\left(T_y \mathbf{a}\right)(x) := \mathbf{a}(x+y), \tag{0.9}$$

and extend this to elements of $\mathcal{F}$ by defining $T_y E := \left\{T_y \mathbf{a} : \mathbf{a} \in E\right\}$.

Except where indicated otherwise, we assume throughout the book that $\mathbb{P}$ is a probability measure on the measurable space $(\Omega, \mathcal{F})$ satisfying the following two important assumptions:

- *Stationarity with respect to $\mathbb{Z}^d$-translations:*

$$\mathbb{P} \circ T_z = \mathbb{P} \quad \text{for every } z \in \mathbb{Z}^d. \tag{0.10}$$

- *Unit range of dependence:*

$$\begin{aligned} &\mathcal{F}_U \text{ and } \mathcal{F}_V \text{ are } \mathbb{P}\text{-independent for every pair } U, V \subseteq \mathbb{R}^d \\ &\text{of Borel subsets satisfying } \operatorname{dist}(U, V) \geqslant 1. \end{aligned} \tag{0.11}$$

We denote the expectation with respect to $\mathbb{P}$ by $\mathbb{E}$. That is, if $X : \Omega \to \mathbb{R}$ is an $\mathcal{F}$-measurable random variable, we write

$$\mathbb{E}[X] := \int_\Omega X(\mathbf{a}) d\mathbb{P}(\mathbf{a}). \tag{0.12}$$

While all random objects we study in this text are functions of $\mathbf{a} \in \Omega$, we do not typically display this dependence explicitly in our notation. We rather use the symbol $\mathbf{a}$ or $\mathbf{a}(x)$ to denote the *canonical coefficient field* with law $\mathbb{P}$.

Examples Satisfying the Assumptions

Perhaps, the simplest way to construct explicit examples satisfying the assumptions of uniform ellipticity (0.6), stationarity (0.10), and (0.11) is by means of a "random checkerboard" structure: we pave the space by unit-sized cubes and color each cube either white or black independently at random. Each color is then associated with a particular value of the diffusivity matrix. More precisely, let $(b(z))_{z \in \mathbb{Z}^d}$ be independent random variables such that for every $z \in \mathbb{Z}^d$,

$$\mathbb{P}[b(z) = 0] = \mathbb{P}[b(z) = 1] = \frac{1}{2},$$

and fix two matrices $\mathbf{a}_0, \mathbf{a}_1$ belonging to the set

$$\left\{ \tilde{\mathbf{a}} \in \mathbb{R}^{d \times d}_{\text{sym}} : \forall \xi \in \mathbb{R}^d, \quad |\xi|^2 \leqslant \xi \cdot \tilde{\mathbf{a}} \xi \leqslant \Lambda |\xi|^2 \right\}. \tag{0.13}$$

We can then define a random field $x \mapsto \mathbf{a}(x)$ satisfying (0.6) and with a law satisfying (0.10) and (0.11) by setting, for every $z \in \mathbb{Z}^d$ and $x \in z + \left[-\frac{1}{2}, \frac{1}{2}\right)^d$,

$$\mathbf{a}(x) = \mathbf{a}_{b(z)}.$$

This example is illustrated in Fig. 1. It can be generalized as follows: we give ourselves a family $(\mathbf{a}(z))_{z \in \mathbb{Z}^d}$ of independent and identically distributed (i.i.d.) random variables taking values in the set (0.13), and then extend the field $z \mapsto \mathbf{a}(z)$ by setting, for every $z \in \mathbb{Z}^d$ and $x \in z + \left[-\frac{1}{2}, \frac{1}{2}\right)^d$,

$$\mathbf{a}(x) := \mathbf{a}(z).$$

Another class of examples can be constructed using homogeneous Poisson point processes. We recall that a Poisson point process on a measurable space $(E, \mathcal{E})$ with (non-atomic, σ-finite) intensity measure μ is a random subset Π of E such that the following properties hold (see also [86]):

- For every measurable set $A \in \mathcal{E}$, the number of points in $\Pi \cap A$, which we denote by $N(A)$, follows a Poisson law of mean $\mu(A)$;
- For every pairwise disjoint measurable sets $A_1, \ldots, A_k \in \mathcal{E}$, the random variables $N(A_1)$, …, $N(A_k)$ are independent.

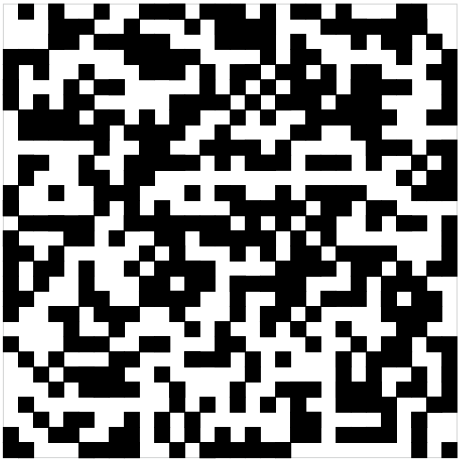

Fig. 1 A piece of a sample of a random checkerboard. The conductivity matrix is equal to $\mathbf{a}_0$ in the black region, and $\mathbf{a}_1$ in the white region

Let Π be a Poisson point process on $\mathbb{R}^d$ with intensity measure given by a multiple of the Lebesgue measure. Fixing two matrices $\mathbf{a}_0, \mathbf{a}_1$ belonging to the set (0.13), we may define a random field $x \mapsto \mathbf{a}(x)$ by setting, for every $x \in \mathbb{R}^d$,

$$\mathbf{a}(x) := \begin{cases} \mathbf{a}_0 & \text{if } \operatorname{dist}(x, \Pi) \leqslant \frac{1}{2}, \\ \mathbf{a}_1 & \text{otherwise.} \end{cases} \tag{0.14}$$

This example is illustrated in Fig. 2. Other similar examples can be constructed by varying the point processes or by replacing balls by different, possibly random, shapes. For instance, choose $\lambda > 0$ and let μ denote a probability measure on $\left[0, \frac{1}{2}\right]$ (the law of the random radius), and let Π be a Poisson point process on $\mathbb{R}^d \times \mathbb{R}$ with intensity measure $\lambda dx \otimes \mu$ (where dx denotes the Lebesgue measure on $\mathbb{R}^d$). We may then define, for every $x \in \mathbb{R}^d$,

$$\mathbf{a}(x) := \begin{cases} \mathbf{a}_0 & \text{if there exists } (z, r) \in \Pi \text{ such that } |x - z| \leqslant r, \\ \mathbf{a}_1 & \text{otherwise.} \end{cases}$$

By varying the construction above, one may replace balls by random shapes, allow the conductivity matrix to take more than two values, and so forth. See Fig. 3 for an example.

A further class of examples can be obtained by defining the coefficient field $x \mapsto \mathbf{a}(x)$ as a local function of a white noise field (we refer to Definition 5.1 and Proposition 5.14 for the definition and construction of white noise). Given a scalar white noise W and a smooth function $\phi \in C_c^\infty(\mathbb{R}^d)$ with support in $B_{1/2}$, a smooth function F from $\mathbb{R}^d$ into the set (0.13), we may define

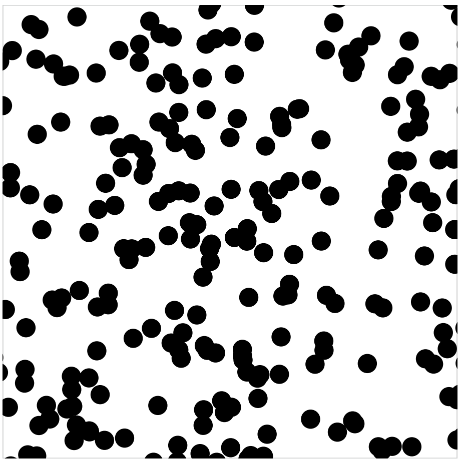

Fig. 2 A sample of the coefficient field defined in (0.14) by the Poisson point cloud. The matrix $\mathbf{a}$ is equal to $\mathbf{a}_0$ in the black region and to $\mathbf{a}_1$ in the white region

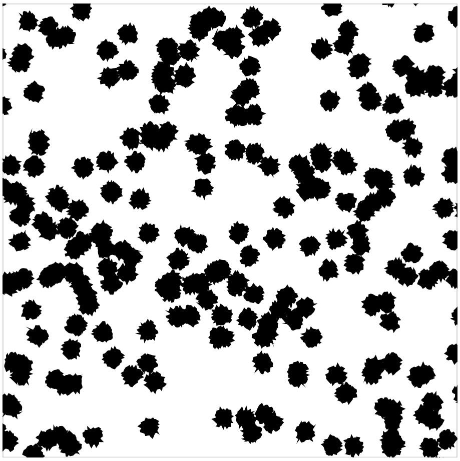

Fig. 3 This coefficient field is sampled from the same distribution as in Fig. 2, except that the balls have been replaced by random shapes

$$\mathbf{a}(x) = F((\mathrm{W} * \phi)(x)). \tag{0.15}$$

See Fig. 4 for a representation of the scalar field $x \mapsto (\mathrm{W} * \phi)(x)$.

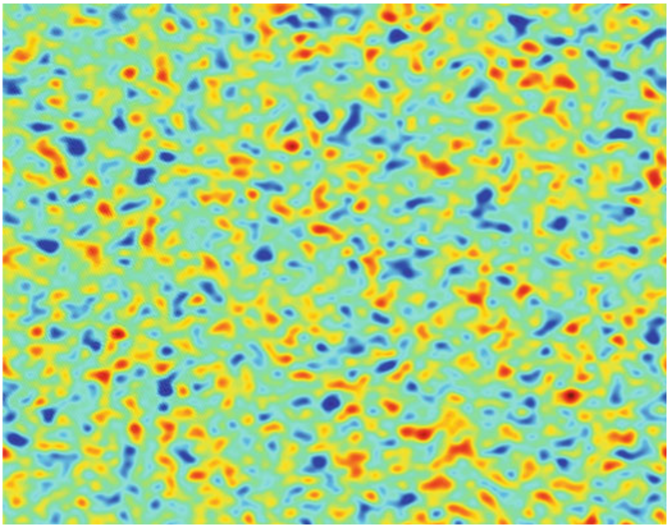

Fig. 4 The figure represents the convolution of white noise with a smooth function of compact support, using a color scale. This scalar field can be used to construct a matrix field $x \mapsto \mathbf{a}(x)$ satisfying our assumptions, see (0.15)

Frequently Asked Questions

Where is the Independence Assumption Used?

The unit range of dependence assumption (0.11) is obviously very important, and to avoid diluting its power we use it sparingly. We list here all the places in the book where it is invoked:

- The proof of Proposition 1.7 (which is made redundant by the following one).
- The proof of Lemma 2.13 (and the generalizations of this lemma appearing in Chaps. 10 and 11). This lemma lies at the heart of the iteration argument in Chap. 2, as it is here that we obtain our first estimate on the correspondence between spatial averages of gradients and fluxes of solutions. Notice that the proof does not use the full strength of the independence assumption; it actually requires only a very weak assumption of correlation decay.
- The last step of the proof of Theorem 2.4 (and the generalizations of this theorem appearing in Chaps. 10 and 11). Here independence is used very strongly to obtain homogenization estimates with optimal stochastic integrability.
- The proof of Proposition 4.12 in Sect. 4.5, where we control the fluctuations of the quantity J_1 inside the bootstrap argument.
- The proof of Proposition 4.27 in Sect. 4.7, where we prove sharp, pointwise-type bounds on the first-order correctors in dimension $d = 2$ (with the correct power of the logarithm).
- In Sect. 5.4, where we prove the central limit theorem for the quantity J_1. This can be considered a refinement of Proposition 4.12.
- In Sect. 9.1, in the proofs of Lemmas 9.7 and 9.10.

In particular, all of the results of Chaps. 2 and 3 are obtained with only two very straightforward applications of independence.

Can the Independence Assumption Be Relaxed?

Yes. One of the advantages of the approach presented here is that *the independence assumption is applied only to sums of local random variables.* Any reasonable decorrelation condition or mixing-type assumption will give estimates regarding the stochastic cancellations of sums of local random variables (indeed, this is essentially a tautology). Therefore, while the statements of the theorems may need to be modified for weaker assumptions (for instance, the strong stochastic integrability results we obtain under a finite range of dependence assumption may have to be weakened), the proofs will only require straightforward adaptations. Since we have only used independence in a handful of places in the text, enumerated above, it is not a daunting task to perform these adaptations. In particular, when the law of the coefficient field satisfies a spectral gap or log-Sobolev assumption (as considered in the series of recent works of Gloria, Otto, and their collaborators), one can easily adapt the arguments presented in this book using concentration inequalities such as [29, Theorem 6.7].

The reason for formalizing the results under the strongest possible mixing assumption (finite range of dependence) rather than attempting to write a very general result is, therefore, not due to a limitation of the arguments. It is simply because we favor clarity of exposition over generality.

Can the Uniform Ellipticity Assumption be Relaxed?

One of the principles of this book is that one should avoid using small-scale or pointwise properties of the solutions or of the equation and focus rather on large-scale, averaged information. We concentrate, especially in the first part of the book, on the energy quantities μ, μ^* , and J_1 which can be thought of as "coarsened coefficients" in analogy to a renormalization scheme (see Remark 2.3). The arguments we use adhere to this philosophy rather strictly. As a result, they are adaptable to situations in which the matrix $\mathbf{a}(x)$ is not necessarily uniformly positive definite, provided we have some quantitative information, for instance, regarding the law of its condition number. This is because such assumptions can be translated into quantitative information about J_1 which suffices to run the renormalization arguments of Chap. 2. Indeed, a demonstration of the robustness of these methods can also be found in [9], which adapted Chaps. 2 and 3 of this manuscript to obtain the first quantitative homogenization results on *supercritical percolation clusters* (a particularly extreme example of a degenerate environment).

Do the Results in This Book Apply to Elliptic Systems?

Since the notation for elliptic systems is a bit distracting, we have decided to use scalar notation. However, throughout most of the book, we use exclusively arguments which also work for systems of equations (satisfying the uniform ellipticity assumption). The only exceptions are the last two sections of Chaps. 8 and 9, where we do use some scalar estimates (the De Giorgi–Nash L^∞ bound and variations) which make it easier to work with Green functions. In particular, we claim that all of the statements and proofs appearing in this book, with the exception of those appearing in those two chapters, can be adapted to the case of elliptic systems with easy and straightforward modifications to the notation.

This Book is Written for Equations in the Continuum. Do the Arguments Apply to Finite Difference Equations on $\mathbb{Z}^d$?

The techniques developed in the book are robust to the underlying structure of the environment on the microscopic scale. What is important is that the "geometry" of the macroscopic medium is like that of $\mathbb{R}^d$ in the sense that certain functional inequalities (such as the Sobolev inequality) are valid, at least on sufficiently large scales. In the case where $\mathbb{R}^d$ is replaced by $\mathbb{Z}^d$, the modifications are relatively straightforward: besides changes to the notation, there is just the slight detail that the boundary of a large cube has a nonzero volume, which creates an additional error term in Chap. 2, causing no harm. If one has a more complicated microstructure like a random graph, such as a supercritical Bernoulli percolation cluster, it is necessary to first establish the "geometric regularity" of the graph in the sense that Sobolev-type inequalities are valid on large scales. The techniques described in this book can then be readily applied: see [9, 42].

Is There a Simple Proof of *Qualitative* Homogenization Somewhere Here?

The arguments in Chap. 1 only need to be slightly modified in order to obtain a more general qualitative homogenization result valid in the case where the unit range of dependence assumption is relaxed to mere *ergodicity*. In other words, in place of (0.11), we assume instead that

$$\text{if}\, A \in \mathcal{F} \,\text{satisfies}\, T_z A = A \,\text{for all}\, z \in \mathbb{Z}^d, \text{then}\, \mathrm{P}[A] \in \{0, 1\}. \tag{0.16}$$

In fact, the only argument that needs to be modified is the proof of Proposition 1.7, since it is the only place in the chapter where independence is used. Moreover, that argument is essentially a proof of the subadditive ergodic theorem in the special case of the unit range of dependence assumption (0.11). In the general ergodic case (0.16), one can simply directly apply the subadditive ergodic theorem (see for instance [3]) to obtain, in place of (1.30), the estimate

$$\mathbb{P}\left[\limsup_{n\to\infty} |\mathbf{a}(\square_n) - \bar{\mathbf{a}}| = 0\right] = 1.$$

The other arguments in that chapter are deterministic and imply that the random variable $\mathcal{E}'(\varepsilon)$ in Theorem 1.17 satisfies $\mathbb{P}[\limsup_{\varepsilon\to 0} \mathcal{E}'(\varepsilon) = 0] = 1$.

What Do We Learn About Reversible Diffusions in Random Environments From the Results in This Book?

A great amount of information about a Markov process can be obtained by studying its infinitesimal generator. Therefore, just as we learn about Brownian motion from properties of harmonic functions (and conversely), the study of divergence-form operators gives us information about the associated diffusion processes.

In fact, any divergence-form elliptic operator is the infinitesimal generator of a reversible diffusion process. Indeed, the De Giorgi–Nash–Aronson estimates recalled in (E.7) and Proposition E.3 can be used together with the classical Kolmogorov extension and continuity theorems (see [25, Theorem 36.2] and [119, Theorem I.2.1]) to define the corresponding stochastic process. Denoting by $\mathbf{P}_x^{\mathbf{a}}$ the probability law of the diffusion process starting from $x \in \mathbb{R}^d$, and by $(X(t))_{t \geqslant 0}$ the canonical process, we have by construction that, for every $\mathbf{a} \in \Omega$, Borel measurable set $A \subseteq \mathbb{R}^d$ and $(t,x) \in (0,\infty) \times \mathbb{R}^d$,

$$\mathbf{P}_x^{\mathbf{a}}[X_t \in A] = \int_A P(t,x,y)\,dy, \tag{0.17}$$

where $P(t,x,y)$ is the parabolic Green function defined in Proposition E.1. The statement

$$\text{for every } x \in \mathbb{R}^d, \qquad t^{\frac{d}{2}} P(t,0,t^{\frac{1}{2}}x) \xrightarrow[t\to\infty]{\text{a.s.}} \bar{P}(1,0,x),$$

where $\overline{P}$ is the parabolic Green function for the homogenized operator, can thus be interpreted as a (quenched) local central limit theorem for the diffusion process. (Note that this also implies convergence in law of the rescaled stochastic process to the Brownian motion with covariance matrix given by $2\bar{\mathbf{a}}$.) Seen in this light, Theorem 8.20 gives us a first quantitative version of this local central limit theorem. The much more precise Theorem 9.11 gives an optimal rate of convergence for this statement, and can thus be interpreted as analogous to the classical Berry–Esseen theorem on the rate of convergence in the central limit theorem for sums of independent random variables (see [117, Theorem 5.5]).

Do the Assumptions Allow for Deterministic, Periodic Coefficient Fields?

Yes, but it would be crazy to use this book as a way to learn about the periodic case, because it is in many ways simpler than the random case. Most of the analysis presented in the first half of the book can be skipped and replaced by the trivial statement that the first-order correctors are periodic functions. The rest of the arguments in the book can be drastically simplified. Readers interested in the quantitative theory for periodic homogenization would be better off with the recent book of Shen [123].

Are the Methods Presented Here Useful for the Study of Homogenization of Other Types of Partial Differential Equations?

Homogenization is closely related to the concept of *averaging* and both are terms which may be used in descriptions of a wide variety of mathematical and physical phenomena. It is impossible to develop a theory of homogenization that would apply to all possible situations or equations, just as it is impossible to develop a regularity theory that would apply to general equations. In this book, we focus solely on the homogenization of elliptic and parabolic equations in *divergence form*, which is the simplest and most physically relevant class of PDEs which generalize the Laplace equation. The reader should expect that the methods and results presented here are useful for equations which are closely related to divergence-form elliptic or parabolic equations. There are, of course, multitudes of important works concerning homogenization for other classes of equations describing a variety of multiscale phenomena which are not described in this book.

For instance, there is a rather mature theory of stochastic homogenization for another class of equations generalizing the Laplace and heat equations, namely, elliptic and parabolic equations in *nondivergence* form (see [115, 89, 132, 34, 33, 17] and the references therein for the qualitative theory and [16, 92, 13] for the more recent quantitative theory). A typical example may take the form

$$-\mathrm{tr}\big(\mathbf{a}(x)\nabla^2 u\big) := -\sum_{i,j=1}^{d} \mathbf{a}_{ij}(x)\partial_{x_i}\partial_{x_j} u = f(x) \quad \text{in } U \subseteq \mathbb{R}^d.$$

Whereas equations in divergence form characterize *reversible* diffusions, equations in nondivergence form describe *balanced* diffusions. While the theory of homogenization for nondivergence form equations shares some high-level ideas and philosophy with the one developed in this book, the two theories are necessarily quite distinct and in particular have no direct implication on each other—even if the coefficients are assumed to be smooth! In fact, the scaling of the optimal quantitative estimates as well as the large-scale regularity theory is quite different. This is to be expected, since the classical regularity theories for equations in divergence and nondivergence form (with rapidly oscillating coefficients) are also distinct.

What About Equations with Locally Stationary Coefficients or with Lower-Order Terms?

More general equations can be considered, such as

$$-\nabla \cdot \left(\mathbf{a}\left(x, \frac{x}{\varepsilon}\right)\nabla u^{\varepsilon}\right) + \mathbf{b}\left(x, \frac{x}{\varepsilon}\right) \cdot \nabla u^{\varepsilon} + c\left(x, \frac{x}{\varepsilon}\right)u^{\varepsilon} = 0. \tag{0.18}$$

In addition to allowing for lower-order terms, the equation above is only “locally stationary” because it allows the coefficients to depend as well on the macroscopic variable x. One could assume, for instance, that, for each x, the field $(\mathbf{a}(x,\cdot), \mathbf{b}(x,\cdot), c(x,\cdot))$ satisfies our usual assumptions of stationarity and finite range dependence, and that the dependence in x is sufficiently regular.

We do not present results for the homogenization of such equations in this book, in an effort to keep things simple and because, in fact, it is relatively routine to handle an equation like (0.18) once one has good control of the first-order correctors $\phi_e(x,\cdot)$ for the coefficient field $\mathbf{a}(x,\cdot)$—that is, for each frozen x and with the lower order terms ignored. Indeed, once the correctors are under control, what is needed is "just" a two-scale expansion computation, similar to what is found in several places in this book and which is essentially no more difficult than in the periodic setting.

Notation

Sets and Euclidean Space

The set of nonnegative integers is denoted by $\mathbb{N} := \{0,1,2,3,\ldots\}$, the set of integers by $\mathbb{Z}$, the set of natural numbers by $\mathbb{N}_* := \mathbb{N}\setminus\{0\}$, the set of rational numbers by $\mathbb{Q}$, and the set of real numbers by $\mathbb{R}$. When we write $\mathbb{R}^m$, we implicitly assume that $m \in \mathbb{N}$. For each $x, y \in \mathbb{R}^m$, the scalar product of x and y is denoted by $x \cdot y$, their tensor product by $x \otimes y$ and the Euclidean norm on $\mathbb{R}^m$ is $|\cdot|$. The canonical basis of $\mathbb{R}^m$ is written as $\{e_1,\ldots,e_m\}$. We let $\mathcal{B}$ denote the Borel σ-algebra on $\mathbb{R}^m$. A *domain* is an open connected subset of $\mathbb{R}^m$. The notions of $C^{k,\alpha}$ *domain* and *Lipschitz domain* are defined in Definition B.1. The boundary of $U \subseteq \mathbb{R}^m$ is denoted by ∂U and its closure by $\overline{U}$. The open ball of radius $r > 0$ centered at $x \in \mathbb{R}^m$ is $B_r(x) := \{y \in \mathbb{R}^m : |x-y| < r\}$. The distance from a point to a set $V \subseteq \mathbb{R}^m$ is written $\operatorname{dist}(x,V) := \inf\{|x-y| : y \in V\}$. For $r > 0$ and $U \subseteq \mathbb{R}^m$, we define

$$U_r := \{x \in U : \operatorname{dist}(x,\partial U) > r\} \quad \text{and} \quad U^r := \{x \in \mathbb{R}^m : \operatorname{dist}(x,U) < r\}. \tag{0.19}$$

For $\lambda > 0$, we set $\lambda U := \{\lambda x : x \in U\}$. If $m,n \in \mathbb{N}$, the set of $m \times n$ matrices with real entries is denoted by $\mathbb{R}^{m\times n}$. We typically denote an element of $\mathbb{R}^{m\times n}$ by a boldfaced latin letter, such as $\mathbf{m}$, and its entries by $(\mathbf{m}_{ij})$. The subset of $\mathbb{R}^{n\times n}$ of symmetric matrices is written $\mathbb{R}^{n\times n}_{\mathrm{sym}}$, and the set of n-by-n skew-symmetric matrices is $\mathbb{R}^{n\times n}_{\mathrm{skew}}$. The identity matrix is denoted Id. If $r,s \in \mathbb{R}$ then we write $r \vee s := \max\{r,s\}$ and $r \wedge s := \min\{r,s\}$. We also define $r_+ := r \vee 0$ and $r_- := -(r \wedge 0)$.

We use triadic cubes throughout the book. For each $m \in \mathbb{N}_0$, we define

$$\square_m := \left(-\frac{1}{2}3^m, \frac{1}{2}3^m\right)^d \subseteq \mathbb{R}^d. \tag{0.20}$$

Observe that, for each $n \in \mathbb{N}_0$ with $n \leqslant m$, the cube $\square_m$ can be partitioned (up to a set of zero Lebesgue measure) into exactly $3^{d(m-n)}$ subcubes which are $\mathbb{Z}^d$-translations of $\square_n$, namely $\left\{z + \square_n : z \in 3^n\mathbb{Z}^d \cap \square_m\right\}$.

Calculus

If $U \subseteq \mathbb{R}^d$ and $f : U \to \mathbb{R}$, we denote the partial derivatives of f by $\partial_{x_i} f$ or simply $\partial_i f$, which unless otherwise indicated, is to be understood in the sense of distributions. A vector field on $U \subseteq \mathbb{R}^d$ is a map $\mathbf{f} : U \to \mathbb{R}^d$. The divergence of a vector field $\mathbf{f}$ is $\nabla \cdot \mathbf{f} = \sum_{i=1}^d \partial_i f_i$, where the (f_i) are the entries of $\mathbf{f}$, i.e., $\mathbf{f} = (f_1, \ldots f_d)$. The gradient of a function $f : U \to \mathbb{R}$ is denoted by $\nabla f = (\partial_1 f, \ldots, \partial_d f)$. The Hessian of f is denoted by $\nabla^2 f := \left(\partial_i \partial_j f\right)_{i,j \in \{1,\ldots,d\}}$ and, more generally, $\nabla^k f$ is the tensor of kth partial derivatives of f, defined by

$$\nabla^k f := (\partial_{i_1} \cdots \partial_{i_k} f)_{i_1,\ldots,i_k \in \{1,\ldots,d\}}.$$

A d-dimensional multi-index is an element of $\mathbb{N}_0^d$. If $\alpha = (\alpha_1, \ldots, \alpha_d)$ is a multi-index, then we define

$$|\alpha| := \sum_{i=1}^d \alpha_i \quad \text{and} \quad \alpha! := \prod_{i=1}^d \alpha_i!,$$

the higher-order partial derivative ∂^α by

$$\partial^\alpha f := \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d} f$$

and the multinomial x^α by

$$x^\alpha := \prod_{i=1}^d x_i^{\alpha_i}.$$

We think of $\nabla^k f$ as a tensor indexed by multi-indices $\alpha \in \mathbb{N}_0^d$ with $|\alpha| = k$, that is, we may write $\nabla^k f = \left\{(\partial^\alpha f)_\alpha : \alpha \in \mathbb{N}_0^d, |\alpha| = k\right\}$. For each $k \in \mathbb{N}_0$, we also denote by $x^{\otimes k}$ the tensor indexed by multi-indices $\alpha \in \mathbb{N}_0^d$ with $|\alpha| = k$ with entries $\left(x^{\otimes k}\right)_\alpha = x^\alpha$. This gives us a compact notation for writing a Taylor series: for instance, we may express

$$\sum_{k=0}^m \sum_{\alpha \in \mathbb{N}_0^d, |\alpha| = k} \frac{1}{\alpha!} \partial^\alpha f(x_0)(x - x_0)^\alpha$$

as

$$\sum_{k=0}^m \frac{1}{k!} \nabla^k f(x_0)(x - x_0)^{\otimes k}. \tag{0.21}$$

Hölder and Lebesgue Spaces

For $k \in \mathbb{N} \cup \{\infty\}$, the set of functions $f : U \to \mathbb{R}$ which are k times continuously differentiable in the classical sense is denoted by $C^k(U)$. We denote by $C_c^k(U)$ the collections of $C^k(U)$ functions with compact support in U. For $k \in \mathbb{N}$ and $\alpha \in (0,1]$, we denote the classical Hölder spaces by $C^{k,\alpha}(U)$, which are the functions $u \in C^k(U)$ for which the norm

$$\|u\|_{C^{k,\alpha}(U)} := \sum_{n=0}^{k} \sup_{x \in U} |\nabla^n u(x)| + \left[\nabla^k u\right]_{C^{0,\alpha}(U)}$$

is finite, where $[\,\cdot\,]_{C^{0,\alpha}(U)}$ is the seminorm defined by

$$[u]_{C^{0,\alpha}(U)} := \sup_{x,y \in U, x \neq y} \frac{|u(x) - u(y)|}{|x - y|^{\alpha}}.$$

For every Borel set $U \in \mathcal{B}$, we denote by $|U|$ the Lebesgue measure of U. For an integrable function $f : U \to \mathbb{R}$, we may denote the integral of f in a compact notation by

$$\int_U f := \int_U f(x)\,dx.$$

For $U \subseteq \mathbb{R}^d$ and $p \in [1, \infty]$, we denote by $L^p(U)$ the Lebesgue space on U with exponent p, that is, the set of measurable functions $f : U \to \mathbb{R}$ satisfying

$$\|f\|_{L^p(U)} := \left(\int_U |f|^p \right)^{\frac{1}{p}} < \infty.$$

The vector space of functions on U which belong to $L^p(V)$ whenever V is bounded and $\overline{V} \subseteq U$ is denoted by $L^p_{\mathrm{loc}}(U)$. If $|U| < \infty$ and $f \in L^1(U)$, then we write

$$\fint_U f := \frac{1}{|U|} \int_U f.$$

The average of a function $f \in L^1(U)$ on U is also sometimes denoted by

$$(f)_U := \fint_U f.$$

To make it easier to keep track of scalings, we very often work with normalized versions of L^p norms: for every $p \in [1, \infty)$ and $f \in L^p(U)$, we set

$$\|f\|_{\underline{L}^p(U)} := \left(\fint_U |f|^p \right)^{\frac{1}{p}} = |U|^{-\frac{1}{p}} \|f\|_{L^p(U)}.$$

For convenience, we may also use the notation $\|f\|_{\underline{L}^\infty(U)} := \|f\|_{L^\infty(U)}$. If X is a Banach space, then $L^p(U;X)$ denotes the set of measurable functions $f : U \to X$ such that $x \mapsto \|f(x)\|_X \in L^p(U)$. We denote the corresponding norm by $\|f\|_{L^p(U;X)}$. By abuse of notation, we will sometimes write $\mathbf{f} \in L^p(U)$ if $\mathbf{f} : U \to \mathbb{R}^m$ is a vector field such that $|\mathbf{f}| \in L^p(U)$ and define $\|\mathbf{f}\|_{\underline{L}^p(U)} := \|\mathbf{f}\|_{\underline{L}^p(U;\mathbb{R}^m)} = \||\mathbf{f}|\|_{\underline{L}^p(U)}$. For $f \in L^p(\mathbb{R}^d)$ and $g \in L^{p'}(\mathbb{R}^d)$ with $\frac{1}{p} + \frac{1}{p'} = 1$, the convolution of f and g is defined as

$$(f * g)(x) := \int_{\mathbb{R}^d} f(x-y) g(y)\, dy.$$

The (essential) *support* of a function $f : U \to \mathbb{R}$ is

$$\operatorname{supp} f := U \setminus \cup \left\{ B_r(x) : x \in \mathbb{R}^d, r > 0, |\{z \in U : f(z) \neq 0\} \cap B_r(x)| = 0 \right\}.$$

Special Functions

For $p \in \mathbb{R}^d$, we denote the affine function with slope p passing through the origin by

$$\ell_p(x) := p \cdot x.$$

Unless otherwise indicated, $\zeta \in C_c^\infty(\mathbb{R}^d)$ denotes the standard mollifier

$$\zeta(x) := \begin{cases} c_d \exp\left(-(1-|x|^2)^{-1}\right) & \text{if} \quad |x| < 1, \\ 0 & \text{if} \quad |x| \geqslant 1, \end{cases} \tag{0.22}$$

with the multiplicative constant c_d chosen so that $\int_{\mathbb{R}^d} \zeta = 1$. We define, for $\delta > 0$,

$$\zeta_\delta(x) := \delta^{-d} \zeta\left(\frac{x}{\delta}\right). \tag{0.23}$$

The standard heat kernel is denoted by

$$\Phi(t,x) := (4\pi t)^{-\frac{d}{2}} \exp\left(-\frac{|x|^2}{4t}\right) \tag{0.24}$$

and we define, for each $z \in \mathbb{R}^d$ and $r > 0$,

$$\Phi_{z,r}(x) := \Phi(r^2, x - z) \quad \text{and} \quad \Phi_r := \Phi_{0,r}.$$

We also denote by $\mathcal{P}_k$ the set of real polynomials on $\mathbb{R}^d$ of order at most k.