DIGITAL NETS AND SEQUENCES

Indispensable for students, invaluable for researchers, this comprehensive treatment of contemporary quasi–Monte Carlo methods, digital nets and sequences, and discrepancy theory starts from scratch with detailed explanations of the basic concepts and then advances to current methods used in research. As deterministic versions of the Monte Carlo method, quasi–Monte Carlo rules have increased in popularity, with many fruitful applications in mathematical practice. These rules require nodes with good uniform distribution properties, and digital nets and sequences in the sense of Niederreiter are known to be excellent candidates. Besides the classical theory, the book contains chapters on reproducing kernel Hilbert spaces and weighted integration, duality theory for digital nets, polynomial lattice rules, the newest constructions by Niederreiter and Xing and many more. The authors present an accessible introduction to the subject based mainly on material taught in undergraduate courses with numerous examples, exercises and illustrations.

DIGITAL NETS AND SEQUENCES

Discrepancy Theory and Quasi–Monte Carlo Integration

JOSEF DICK
University of New South Wales, Sydney

FRIEDRICH PILLICHSHAMMER
Johannes Kepler Universität Linz

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521191593

First published 2010

Printed in the United Kingdom at the University Press, Cambridge

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

ISBN 978-0-521-19159-3 Hardback

To Jingli and Gisi

Contents

Preface

The theory of digital nets and sequences has its roots in uniform distribution modulo one and in numerical integration using quasi–Monte Carlo (QMC) rules. The subject can be traced back to several influential works: the notion of uniform distribution to a classical paper by Weyl [265]; the Koksma–Hlawka inequality, which forms the starting point for analysing QMC methods for numerical integration, to Koksma [121] in the one-dimensional case and to Hlawka [111] in arbitrary dimension. Explicit constructions of digital sequences were first introduced by Sobol′ [253], followed by Faure [68] and Niederreiter [173]. A general principle of these constructions was introduced by Niederreiter in [172], which now forms one of the essential pillars of QMC integration and of this book. These early results are well summarised in references [61, 114, 130, 171 and 177], where much more information on the history and on earlier discoveries can be found.

Since then, numerical integration based on QMC has been developed into a comprehensive theory with many new facets. The introduction of reproducing kernel Hilbert spaces by Hickernell [101] furnished many Koksma–Hlawka-type inequalities. The worst-case integration error can be expressed directly in terms of a reproducing kernel, a function which, together with a uniquely defined inner product, describes a Hilbert space of functions.

Contrary to earlier suppositions, QMC methods are now used for the numerical integration of functions in hundreds or even thousands of dimensions. The success of this approach has been described by Sloan and Woźniakowski in [249], where the concept of weighted spaces was introduced. These weighted spaces nowadays permeate the literature on high-dimensional numerical integration. The result was a weighted Koksma–Hlawka inequality which yields weighted quality measures (called discrepancies) of the quadrature points and the need for the construction of point sets which are of high quality with respect to this new criterion. This led to computer search algorithms for suitable quadrature points which were first developed for lattice rules [246, 247] and subsequently extended to polynomial lattice rules [45].

The construction of low-discrepancy point sets and sequences has also undergone dramatic improvements. The constructions of Sobol′ [253], Faure [68] and Niederreiter [173] have been developed into the prevailing notion of (digital) (t, m, s)-nets and (t, s)-sequences. The problem of asymptotically optimal constructions in the context of this theory (i.e. which minimise the quality parameter t) have been developed by Niederreiter and Xing in [191, 267], with several subsequent extensions. From a theoretical perspective, the development of a duality theory for digital nets is interesting, see [189], which gives a general framework for the theory of digital nets.

Another development has been a partial merging of Monte Carlo (MC) methods, where the quadrature points are chosen purely at random, with QMC. The aim here is to introduce a random element into the construction of low-discrepancy points that, on the one hand, preserves the distribution properties and is, at the same time, sufficiently random to yield an unbiased estimator (and which also has further useful properties). Such a method, called scrambling, has been introduced by Owen [206], and was first analysed in [207, 209]. As a bonus, one can obtain an improved rate of convergence of $O(N^{-3/2}(\log N)^c)$ (for some $c > 0$) using this randomisation.

The topic of improved rates of convergence was further developed first in [104] for lattice rules, and then in [27] for polynomial lattice rules, using a random shift and the tent transformation. This method achieves convergence rates of $O(N^{-2}(\log N)^c)$ (for some $c > 0$). The quadrature points which can be used in this method can be found by computer search.

A general theory of higher order digital nets and sequences has been developed in [35] for periodic functions, and in [36] for the general case. There, the convergence rate is of $O(N^{-\alpha}(\log N)^c)$ (for some $c > 0$), with $\alpha > 1$ arbitrarily large for sufficiently smooth functions.

A breakthrough concerning the classical problem of finding an explicit construction of point sets which achieve the optimal rate of convergence of the L_2-discrepancy came from Chen and Skriganov [22]. This problem goes back to the lower bound on the L_2-discrepancy by Roth [228].

The aim of this work is to describe these achievements in the areas of QMC methods and uniform distribution. The choice and presentation of the topics is naturally biased towards our, the authors, interests and expertise. Another consideration for such choice of topics concerns the monographs already available, many of which are cited throughout the book.

In order to give a consistent and comprehensive treatment of the subject, we use Walsh series analysis throughout the book. In a broader context this has already featured in [130, 170] and in the context of analysing digital nets in [133, 148].

Some authors, especially those concerned with the analysis of the mean-square worst-case error of scrambled nets, prefer to use Haar wavelets, which were also used, for instance, by Sobol′ [252, 253].

In the analysis of scrambled nets, no disadvantage seems to arise from replacing Haar functions with Walsh functions. The locality of Haar functions is offset by the locality of the Walsh–Dirichlet kernel. As illustration, Owen's description of a nested Analysis of Variance (ANOVA) decomposition [207] can also be neatly described using the Walsh–Dirichlet kernel; see Section 13.2. For where Walsh functions are seen to be of considerable advantage, see Chapter 14. The Walsh coefficients of smooth functions exhibit a certain decay which is an essential ingredient in the theory of higher order digital nets and sequences. This property is not shared in the same way by Haar coefficients of smooth functions. Furthermore, the construction of point sets with optimal L_2 discrepancy has its origin in the Walsh series expansion of the characteristic function $\chi_{[0,x)}$. This makes Walsh functions more suited to our endeavour than Haar functions. However, this does not mean that this is always the case; in future work, researchers should consider such a choice on a case-by-case basis.

The aim of this book is to provide an introduction to the topics described above as well as to some others. Parts of a theory which has already appeared elsewhere are repeated here in order to make the monograph as self-contained as possible. This is complemented by two appendices, one on Walsh functions and one on algebraic function fields. The latter is the underlying basis for the constructions of digital nets and sequences by Niederreiter, Xing and Özbudak described in Chapter 8.

The text is aimed at undergraduate students in mathematics. The exercises at the end of each chapter make it suitable for an undergraduate or graduate course on the topic of this book or parts thereof. Such a course may be useful for students of science, engineering or finance, where QMC methods find their applications. We hope that it may prove useful for our colleagues as a reference book and an inspiration for future work. We also hope for an advancement in the area in the next few decades akin to that which we have seen in the past.

Acknowledgements

The germ of this book goes back many years now to a handwritten manuscript by Gerhard Larcher that was the basis for the first author's master's thesis under Gerhard's supervision and which now forms the main part of Chapters 4 and 5. This manuscript was in fact the first comprehensive introduction to the topic for the authors. For this and other contributions, we are immensely grateful. Thank you, Gerhard!

The text has also greatly benefited from valuable comments and suggestions by several colleagues which we would like to mention here: Jan Baldeaux, Johann Brauchart, Henri Faure, Michael Gnewuch, Stefan Heinrich, Peter Hellekalek, Roswitha Hofer, Stephen Joe, Peter Kritzer, Gerhard Larcher, Gunther Leobacher, Harald Niederreiter, Erich Novak, Art Owen, Gottlieb Pirsic, Wolfgang Ch. Schmid, Ian Sloan, and Henryk Woźniakowski.

It is a pleasure to thank the staff at Cambridge University Press for their professionalism in the production of the book, in particular Clare Dennison, Cheryl Gibson, Diana Gillooly, Abigail Jones and David Tranah.

Finally, we are greatly indebted to our families without whom this book would not exist.

Josef Dick
Friedrich Pillichshammer

Notation

Note: In the following, we list only symbols that are used in a global context.

Some specific sets and numbers

$\mathbb{C}$	Complex numbers.
$\mathbb{F}_b$	Finite field with b elements for a prime power b (if b is a prime, then we identify $\mathbb{F}_b$ with $\mathbb{Z}_b$). The elements of $\mathbb{F}_b$ (for b not a prime) are sometimes denoted by $\overline{0}, \overline{1}, \ldots, \overline{b-1}$.
$\mathbb{F}_b[x]$, $\mathbb{Z}_b[x]$	Set of polynomials over $\mathbb{F}_b$ or $\mathbb{Z}_b$.
$\mathbb{F}_b((x^{-1}))$, $\mathbb{Z}_b((x^{-1}))$	Field of formal Laurent series over $\mathbb{F}_b$ or $\mathbb{Z}_b$.
$G_{b,m}$	$G_{b,m} = \{q \in \mathbb{F}_b[x] : \deg(q) < m\}$.
i	$\mathrm{i} = \sqrt{-1}$.
$\mathcal{I}_s$	Index set $\{1, \ldots, s\}$.
$\mathbb{N}$	Positive integers.
$\mathbb{N}_0$	Non-negative integers.
$\mathcal{P}$	Finite point set in $[0, 1)^s$ (interpreted in the sense of the combinatorial notion of 'multiset', i.e. a set in which the multiplicity of elements matters).
$\mathcal{P}_{\mathfrak{u}}$	Point set in $[0, 1)^{\lvert\mathfrak{u}\rvert}$ consisting of the points from $\mathcal{P}$ projected to the components given by $\mathfrak{u} \subseteq \mathcal{I}_s$.
$\mathbb{R}$	Real numbers.
$\mathcal{S}$	Infinite sequence in $[0, 1)^s$.
$\mathfrak{u}, \mathfrak{v}, \ldots$	Subsets of $\mathcal{I}_s$.
ω_b	$\omega_b = \mathrm{e}^{2\pi\mathrm{i}/b}$.
$\lvert X\rvert$	Cardinality of a set X.
X^m	The m-fold Cartesian product of a set X.

$(X^m)^\top$	The set of m-dimensional column vectors over X.
$\mathbb{Z}$	Integers.
$\mathbb{Z}_b$	Residue class ring modulo b (we identify $\mathbb{Z}_b$ with $\{0, \ldots, b-1\}$ with addition and multiplication modulo b).
$\boldsymbol{\gamma}$	Set of non-negative weights, i.e. $\boldsymbol{\gamma} = \{\gamma_{\mathfrak{u}} : \mathfrak{u} \subseteq \mathcal{I}_s\}$. In the case of product weights, $\boldsymbol{\gamma} = (\gamma_i)_{i \ge 1}$ is understood as the sequence of one-dimensional weights. In this case we set $\gamma_{\mathfrak{u}} = \prod_{i \in \mathfrak{u}} \gamma_i$.

Vectors and matrices

$\mathbf{a}, \mathbf{b}, \mathbf{c}, \ldots, \mathbf{x}, \mathbf{y}, \mathbf{z}$	Row vectors over $\mathbb{F}_b$ or $\mathbb{Z}_b$.
$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \ldots, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$	Row vectors over $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$ or $\mathbb{R}$.
$\mathbf{a}^\top, \mathbf{b}^\top, \ldots$	Transpose of a vector $\mathbf{a}, \mathbf{b}, \ldots$ in $\mathbb{F}_b$ or $\mathbb{Z}_b$.
$A, B, C, D, \ldots$	$m \times m$ or $\mathbb{N} \times \mathbb{N}$ matrices over $\mathbb{F}_b$.
$A^\top$	Transpose of the matrix A.
$C^{(m)}$	Left upper $m \times m$ sub-matrix of a matrix C.
$C^{(m \times n)}$	Left upper $m \times n$ sub-matrix of a matrix C.
$\boldsymbol{x}_{\mathfrak{u}}$	For an s-dimensional vector $\boldsymbol{x} = (x_1, \ldots, x_s)$ and for $\mathfrak{u} \subseteq \mathcal{I}_s$ the $\lvert\mathfrak{u}\rvert$-dimensional vector consisting of the components of $\boldsymbol{x}$ whose index belongs to $\mathfrak{u}$, i.e. $\boldsymbol{x}_{\mathfrak{u}} = (x_i)_{i \in \mathfrak{u}}$. For example, for $\boldsymbol{x} = (\frac{1}{10}, \frac{1}{3}, \frac{1}{5}, \frac{1}{4}, \frac{1}{8}) \in [0,1)^5$ and $\mathfrak{u} = \{2, 3, 5\}$, we have $\boldsymbol{x}_{\mathfrak{u}} = (\frac{1}{3}, \frac{1}{5}, \frac{1}{8})$.
$(\boldsymbol{x}_{\mathfrak{u}}, \boldsymbol{w})$	For $\boldsymbol{w} = (w_1, \ldots, w_s)$, the vector whose ith component is x_i if $i \in \mathfrak{u}$ and w_i if $i \notin \mathfrak{u}$.
$\boldsymbol{x} \cdot \boldsymbol{y}$ (or $\mathbf{x} \cdot \mathbf{y}$)	Usual inner product of the two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ (or $\mathbf{x}$ and $\mathbf{y}$, respectively).
$(\boldsymbol{x}_{\mathfrak{u}}, \mathbf{0})$	For an s-dimensional vector $\boldsymbol{x} = (x_1, \ldots, x_s)$ and for $\mathfrak{u} \subseteq \mathcal{I}_s$ the s-dimensional vector whose ith component is x_i if $i \in \mathfrak{u}$ and 0 if $i \notin \mathfrak{u}$. For example, for $\boldsymbol{x}$ and $\mathfrak{u}$ as above, we have $(\boldsymbol{x}_{\mathfrak{u}}, \mathbf{0}) = (0, \frac{1}{3}, \frac{1}{5}, 0, \frac{1}{8})$.
$(\boldsymbol{x}_{\mathfrak{u}}, \mathbf{1})$	Like $(\boldsymbol{x}_{\mathfrak{u}}, \mathbf{0})$ with zero replaced by one.

Some specific functions

$\overline{a}$	Complex conjugate of a complex number a.
$A(J, N, \mathcal{P})$	For a $\mathcal{P} = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\}$, the number of indices n, $0 \le n < N$, for which the point $\boldsymbol{x}_n$ belongs to J.
$A(J, N, \mathcal{S})$	For $\mathcal{S} = (\boldsymbol{x}_n)_{n \ge 0}$, the number of indices n, $0 \le n < N$, for which the point $\boldsymbol{x}_n$ belongs to J.

${}_b\mathrm{wal}_k$	kth b-adic Walsh function (see Definition A.1).
B_k	kth Bernoulli polynomial.
$d \mid n$, $d \nmid n$	d divides n (d does not divide n).
D_N	Extreme discrepancy (see Definition 3.13).
D_N^*	Star discrepancy (see Definitions 2.2 and 2.14).
$D_{N,\boldsymbol{\gamma}}^*$	Weighted star discrepancy (see Definition 3.59).
$\mathbb{E}$	Expectation.
$I(f)$	Integral of the function f over the s-dimensional unit-cube with respect to the s-dimensional Lebesgue measure, i.e. $I(f) = \int_{[0,1]^s} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$.
$\log x$	Natural logarithm of x.
$\log_b x$	Base b logarithm of x.
$L_{q,N}$	L_q-discrepancy (see Definition 3.19).
$L_{q,N,\boldsymbol{\gamma}}$	Weighted L_q-discrepancy (see Definition 3.59).
$O(f(x))$	For $f, g : \mathbb{R} \to \mathbb{R}$, $f \geq 0$, $g(x) = O(f(x))$ for $x \to a$ if there exist $C, \delta > 0$ such that $\lvert g(x)\rvert \leq Cf(x)$ for all x with $\lvert x - a\rvert < \delta$ (or $x > \delta$ if $a = \infty$).
Prob	Probability.
$Q_N(f)$	Quasi–Monte Carlo (QMC) rule for f and an N-element point set $\mathcal{P} = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\}$, i.e. $Q_N(f) = \frac{1}{N}\sum_{n=0}^{N-1} f(\boldsymbol{x}_n)$.
$\mathrm{tr}_m(\mathbf{k})$	$\mathrm{tr}_m(\mathbf{k}) = (\kappa_0, \ldots, \kappa_{m-1})^\top$ for $k \in \mathbb{N}_0$ with b-adic expansion $k = \sum_{j\geq 0} \kappa_j b^j$.
$\mathrm{tr}_m(k)$	$\mathrm{tr}_m(k) = \kappa_0 + \kappa_1 b + \cdots + \kappa_{m-1} b^{m-1}$ for $k \in \mathbb{N}_0$ with b-adic expansion $k = \sum_{j\geq 0} \kappa_j b^j$.
Var	Variance.
$\{x\}$	Fractional part of a real number x.
$\lfloor x \rfloor$	Integer part of a non-negative real number x, i.e. $\lfloor x \rfloor = x - \{x\}$.
$\lceil x \rceil$	The smallest integer larger than or equal to x.
$(x)_+$	$x_+ = \max(x, 0)$.
$\lvert\boldsymbol{x}\rvert_1$	L_1-norm; $\lvert\boldsymbol{x}\rvert_1 = \lvert x_1\rvert + \cdots + \lvert x_s\rvert$ if $\boldsymbol{x} = (x_1, \ldots, x_s)$.
$\lvert\boldsymbol{x}\rvert_\infty$	Maximum norm; $\lvert\boldsymbol{x}\rvert_\infty = \max_{1\leq i\leq s} \lvert x_i\rvert$ if $\boldsymbol{x} = (x_1, \ldots, x_s)$.
λ_s	s-dimensional Lebesgue measure (for $s = 1$ simply λ).
$\pi_m(\mathbf{c})$	Projection of $\mathbf{c} \in \mathbb{F}_b^{\mathbb{N}} = \{(c_1, c_2, \ldots) : c_1, c_2, \ldots \in \mathbb{F}_b\}$ onto its first m components.
φ	Bijection from $\{0, \ldots, b-1\} \to \mathbb{F}_b$.
φ_b	b-adic radical inverse function (see Definition 3.10).
φ^{-1}	Inverse of the bijection $\varphi : \{0, \ldots, b-1\} \to \mathbb{F}_b$.
$\chi_J(x)$	Characteristic function of a set J, i.e. $\chi_J(x) = 1$ if $x \in J$ and $\chi_J(x) = 0$ if $x \notin J$.