London Mathematical Society
Lecture Note Series 205

Ergodic Theory and its Connections with Harmonic Analysis

Edited by
Karl E. Petersen
and
Ibrahim A. Salama

CAMBRIDGE UNIVERSITY PRESS

CAMBRIDGE

more information – www.cambridge.org/9780521459990

This page intentionally left blank

LONDON MATHEMATICAL SOCIETY LECTURE NOTE SERIES

Managing Editor: Professor J.W.S. Cassels, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, England

The books in the series listed below are available from booksellers, or, in case of difficulty, from Cambridge University Press.

34 Representation theory of Lie groups, M.F. ATIYAH *et al*
36 Homological group theory, C.T.C. WALL (ed)
39 Affine sets and affine groups, D.G. NORTHCOTT
46 p-adic analysis: a short course on recent work, N. KOBLITZ
50 Commutator calculus and groups of homotopy classes, H.J. BAUES
59 Applicable differential geometry, M. CRAMPIN & F.A.E. PIRANI
66 Several complex variables and complex manifolds II, M.J. FIELD
69 Representation theory, I.M. GELFAND *et al*
74 Symmetric designs: an algebraic approach, E.S. LANDER
76 Spectral theory of linear differential operators and comparison algebras, H.O. CORDES
77 Isolated singular points on complete intersections, E.J.N. LOOIJENGA
79 Probability, statistics and analysis, J.F.C. KINGMAN & G.E.H. REUTER (eds)
83 Homogeneous structures on Riemannian manifolds, F. TRICERRI & L. VANHECKE
86 Topological topics, I.M. JAMES (ed)
87 Surveys in set theory, A.R.D. MATHIAS (ed)
88 FPF ring theory, C. FAITH & S. PAGE
89 An F-space sampler, N.J. KALTON, N.T. PECK & J.W. ROBERTS
90 Polytopes and symmetry, S.A. ROBERTSON
92 Representation of rings over skew fields, A.H. SCHOFIELD
93 Aspects of topology, I.M. JAMES & E.H. KRONHEIMER (eds)
94 Representations of general linear groups, G.D. JAMES
95 Low-dimensional topology 1982, R.A. FENN (ed)
96 Diophantine equations over function fields, R.C. MASON
97 Varieties of constructive mathematics, D.S. BRIDGES & F. RICHMAN
98 Localization in Noetherian rings, A.V. JATEGAONKAR
99 Methods of differential geometry in algebraic topology, M. KAROUBI & C. LERUSTE
100 Stopping time techniques for analysts and probabilists, L. EGGHE
101 Groups and geometry, ROGER C. LYNDON
103 Surveys in combinatorics 1985, I. ANDERSON (ed)
104 Elliptic structures on 3-manifolds, C.B. THOMAS
105 A local spectral theory for closed operators, I. ERDELYI & WANG SHENGWANG
106 Syzygies, E.G. EVANS & P. GRIFFITH
107 Compactification of Siegel moduli schemes, C-L. CHAI
108 Some topics in graph theory, H.P. YAP
109 Diophantine analysis, J. LOXTON & A. VAN DER POORTEN (eds)
110 An introduction to surreal numbers, H. GONSHOR
113 Lectures on the asymptotic theory of ideals, D. REES
114 Lectures on Bochner-Riesz means, K.M. DAVIS & Y-C. CHANG
115 An introduction to independence for analysts, H.G. DALES & W.H. WOODIN
116 Representations of algebras, P.J. WEBB (ed)
117 Homotopy theory, E. REES & J.D.S. JONES (eds)
118 Skew linear groups, M. SHIRVANI & B. WEHRFRITZ
119 Triangulated categories in the representation theory of finite-dimensional algebras, D. HAPPEL
121 Proceedings of *Groups - St Andrews 1985*, E. ROBERTSON & C. CAMPBELL (eds)
122 Non-classical continuum mechanics, R.J. KNOPS & A.A. LACEY (eds)
125 Commutator theory for congruence modular varieties, R. FREESE & R. MCKENZIE
126 Van der Corput's method of exponential sums, S.W. GRAHAM & G. KOLESNIK
127 New directions in dynamical systems, T.J. BEDFORD & J.W. SWIFT (eds)
128 Descriptive set theory and the structure of sets of uniqueness, A.S. KECHRIS & A. LOUVEAU
129 The subgroup structure of the finite classical groups, P.B. KLEIDMAN & M.W.LIEBECK
130 Model theory and modules, M. PREST
131 Algebraic, extremal & metric combinatorics, M-M. DEZA, P. FRANKL & I.G. ROSENBERG (eds)
132 Whitehead groups of finite groups, ROBERT OLIVER
133 Linear algebraic monoids, MOHAN S. PUTCHA
134 Number theory and dynamical systems, M. DODSON & J. VICKERS (eds)
135 Operator algebras and applications, 1, D. EVANS & M. TAKESAKI (eds)
136 Operator algebras and applications, 2, D. EVANS & M. TAKESAKI (eds)
137 Analysis at Urbana, I, E. BERKSON, T. PECK, & J. UHL (eds)
138 Analysis at Urbana, II, E. BERKSON, T. PECK, & J. UHL (eds)
139 Advances in homotopy theory, S. SALAMON, B. STEER & W. SUTHERLAND (eds)
140 Geometric aspects of Banach spaces, E.M. PEINADOR and A. RODES (eds)

141 Surveys in combinatorics 1989, J. SIEMONS (ed)
142 The geometry of jet bundles, D.J. SAUNDERS
143 The ergodic theory of discrete groups, PETER J. NICHOLLS
144 Introduction to uniform spaces, I.M. JAMES
145 Homological questions in local algebra, JAN R. STROOKER
146 Cohen-Macaulay modules over Cohen-Macaulay rings, Y. YOSHINO
147 Continuous and discrete modules, S.H. MOHAMED & B.J. MÜLLER
148 Helices and vector bundles, A.N. RUDAKOV et al
149 Solitons, nonlinear evolution equations and inverse scattering, M.A. ABLOWITZ & P.A. CLARKSON
150 Geometry of low-dimensional manifolds 1, S. DONALDSON & C.B. THOMAS (eds)
151 Geometry of low-dimensional manifolds 2, S. DONALDSON & C.B. THOMAS (eds)
152 Oligomorphic permutation groups, P. CAMERON
153 L-functions and arithmetic, J. COATES & M.J. TAYLOR (eds)
154 Number theory and cryptography, J. LOXTON (ed)
155 Classification theories of polarized varieties, TAKAO FUJITA
156 Twistors in mathematics and physics, T.N. BAILEY & R.J. BASTON (eds)
157 Analytic pro-*p* groups, J.D. DIXON, M.P.F. DU SAUTOY, A. MANN & D. SEGAL
158 Geometry of Banach spaces, P.F.X. MÜLLER & W. SCHACHERMAYER (eds)
159 Groups St Andrews 1989 Volume 1, C.M. CAMPBELL & E.F. ROBERTSON (eds)
160 Groups St Andrews 1989 Volume 2, C.M. CAMPBELL & E.F. ROBERTSON (eds)
161 Lectures on block theory, BURKHARD KÜLSHAMMER
162 Harmonic analysis and representation theory for groups acting on homogeneous trees, A. FIGA-TALAMANCA & C. NEBBIA
163 Topics in varieties of group representations, S.M. VOVSI
164 Quasi-symmetric designs, M.S. SHRIKANDE & S.S. SANE
165 Groups, combinatorics & geometry, M.W. LIEBECK & J. SAXL (eds)
166 Surveys in combinatorics, 1991, A.D. KEEDWELL (ed)
167 Stochastic analysis, M.T. BARLOW & N.H. BINGHAM (eds)
168 Representations of algebras, H. TACHIKAWA & S. BRENNER (eds)
169 Boolean function complexity, M.S. PATERSON (ed)
170 Manifolds with singularities and the Adams-Novikov spectral sequence, B. BOTVINNIK
171 Squares, A.R. RAJWADE
172 Algebraic varieties, GEORGE R. KEMPF
173 Discrete groups and geometry, W.J. HARVEY & C. MACLACHLAN (eds)
174 Lectures on mechanics, J.E. MARSDEN
175 Adams memorial symposium on algebraic topology 1, N. RAY & G. WALKER (eds)
176 Adams memorial symposium on algebraic topology 2, N. RAY & G. WALKER (eds)
177 Applications of categories in computer science, M. FOURMAN, P. JOHNSTONE, & A. PITTS (eds)
178 Lower K- and L-theory, A. RANICKI
179 Complex projective geometry, G. ELLINGSRUD *et al*
180 Lectures on ergodic theory and Pesin theory on compact manifolds, M. POLLICOTT
181 Geometric group theory I, G.A. NIBLO & M.A. ROLLIER (eds)
182 Geometric group theory II, G.A. NIBLO & M.A. ROLLIER (eds)
183 Shintani zeta functions, A. YUKIE
184 Arithmetical functions, W. SCHWARZ & J. SPILKER
185 Representations of solvable groups, O. MANZ & T.R. WOLF
186 Complexity: knots, colourings and counting, D.J.A. WELSH
187 Surveys in combinatorics, 1993, K. WALKER (ed)
188 Local analysis for the odd order theorem, H. BENDER & G. GLAUBERMAN
189 Locally presentable and accessible categories, J. ADAMEK & J. ROSICKY
190 Polynomial invariants of finite groups, D.J. BENSON
191 Finite geometry and combinatorics, F. DE CLERK *et al*
192 Symplectic geometry, D. SALAMON (ed)
193 Computer algebra and differential equations, E. TOURNIER (ed)
194 Independent random variables and rearrangement invariant spaces, M. BRAVERMAN
195 Arithmetic of blowup algebras, WOLMER VASCONCELOS
196 Microlocal analysis for differential operators, A. GRIGIS & J. SJÖSTRAND
197 Two-dimensional homotopy and combinatorial group theory, C. HOG-ANGELONI, W. METZLER & A.J. SIERADSKI (eds)
198 The algebraic characterization of geometric 4-manifolds, J.A. HILLMAN
199 Invariant potential theory in the unit ball of C^n, MANFRED STOLL
200 The Grothendieck theory of dessins d'enfant, L. SCHNEPS (ed)
201 Singularities, JEAN-PAUL BRASSELET (ed)
202 The technique of pseudodifferential operators, H.O. CORDES
203 Hochschild cohomology, A. SINCLAIR & R. SMITH
204 Combinatorial and geometric group theory, A.J. DUNCAN, N.D. GILBERT, J. HOWIE (eds)
205 Ergodic theory and its connections with harmonic analysis, K.E. PETERSEN & I. SALAMA (eds)
207 Groups of Lie type and their geometries, W.M. KANTOR & L. DI MARTINO (eds)

London Mathematical Society Lecture Note Series. 205

Ergodic Theory and its Connections with Harmonic Analysis

Proceedings of the 1993 Alexandria Conference

Edited by

Karl E. Petersen
University of North Carolina

and

Ibrahim A. Salama
North Carolina Central University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

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/9780521459990

© Cambridge University Press 1995

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without the written
permission of Cambridge University Press.

First published 1995

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

ISBN 978-0-521-45999-0 paperback

Transferred to digital printing 2007

Cambridge University Press has no responsibility for the persistence or
accuracy of email addresses referred to in this publication.

CONTENTS

PREFACE

Ergodic theory is a crossroads for many branches of mathematics and science, from which it has drawn problems, ideas, and methods, and in which it has found applications.

In recent years the interaction of ergodic theory with several kinds of harmonic analysis has been especially evident: in the Fourier method of proving almost everywhere convergence theorems introduced by Bourgain and developed by Rosenblatt, Wierdl, and others; in the application of the real-variable harmonic analysis of the Stein school to ergodic theory by, for example, Bellow, Boivin, Deniel, Derriennic, Jones, and Rosenblatt; in rigidity theory, where Katok, Margulis, Mostow, Ratner, Spatzier, Zimmer, and others blend ergodic theory with dynamics, noncommutative harmonic analysis, and geometry; in the study of singular measures, spectral properties, and subgroups by Aaronson, Host, Méla, Nadkarni, and Parreau, which has furthered the analysis of nonsingular transformations and has led to progress on old problems like the higher-order mixing conjecture; in the structure theory of single transformations, where concepts like rank, joinings, and approximation are helping workers such as del Junco, King, Lemanczyk, Rudolph, and Thouvenot to classify systems, to understand better the important family of Gaussian processes, and to explore the connections among spectral and dynamical properties; in the combinatorics of adic transformations, introduced by Vershik and pursued by Herman, Kerov, Livshitz, Putnam, and Skau, which makes visible the connections between some invariants of dynamical systems and certain C^*-algebras, especially those associated with group representations; and in the applications of ergodic theory to combinatorial number theory and Diophantine approximation discovered by Furstenberg and developed also by Bergelson, Hindman, Katznelson, and Weiss, which rely in an essential way on techniques from harmonic analysis and spectral theory.

Egypt is a crossroads of many cultures and peoples, convenient to the countries of the Middle East, Africa, and Europe, and roughly equidistant from Asia and America. Our conference, focussed on the interaction of ergodic theory with harmonic analysis, found a natural site in Alexandria, the home of Euclid, Diophantus, Hypatia, Ptolemy, ... (some of whom we could try to claim as early ergodic theorists—Ptolemy, in his celestial mechanics, perhaps even as someone who applied harmonic analysis to dynamics). While no attempt was made to cover the area completely, some important sectors were described in survey talks, and the research talks reported interesting new developments.

Part I of these Proceedings contains three survey articles expanding presentations given at the conference. Each gives an up-to-date description of an area of lively interaction of ergodic theory and harmonic analysis, laying out the background of the field and the sources of its problems, the most important and most interesting recent results, and the current lines of development and outstanding open questions. These papers—by Rosenblatt and Wierdl on Fourier methods in almost everywhere convergence, by Spatzier on rigidity, and by Thouvenot on joinings—should provide convenient starting points for researchers beginning work in these areas.

Part II is a collection of refereed research papers presenting new results on questions related to the theme of the conference. Some of these papers were presented at the meeting, others were contributed later. Two of them (Lesigne, Rudolph) concern the problems in noncommutative harmonic analysis that emerge in trying to understand the nonlinear ergodic averages arising from Furstenberg's diagonal approach to Szemerédi's Theorem, and three others (Forrest, Hendrick, and McCutcheon) treat further developments in this dynamical multiple-recurrence theory. The remainder deal with problems in almost everywhere convergence and a variety of other topics in dynamics.

It is a pleasure to thank the many people and institutions who contributed to the success of the conference and to the preparation of these Proceedings. The National Science Foundation and Institute of Statistics of Cairo University provided financial support. Special thanks are due to Professor Ahmed E. Sarhan and Professor Mahmoud Riad for their hospitality and handling of the official arrangements in Egypt. We also thank Professor Mounir Morsy for welcoming us to his Department of Mathematics at Ain Shams University. Further, we are indebted to the anonymous referees, who improved the papers presented here beyond their original versions, to the many excellent typists who produced the TEX files, and to Lauren Cowles and David Tranah, who smoothly handled the publishing.

Karl Petersen and Ibrahim Salama
Chapel Hill, N.C.
July 11, 1994

TEX and AMS-TEX are trademarks of the American Mathematical Society.

This book was prepared in AMS-TEX by the author.

PART I

SURVEY ARTICLES

POINTWISE ERGODIC THEOREMS VIA HARMONIC ANALYSIS

JOSEPH M. ROSENBLATT AND MÁTÉ WIERDL

CONTENTS

INTRODUCTION

Historical remarks

It has been eighty-five years since Bohl [1909], Sierpiński [1910] and Weyl [1910] proved the now famous equidistribution theorem: if α is an irrational number then the sequence $\alpha, 2\alpha, 3\alpha \dots$ is uniformly distributed mod 1. This means that for each subinterval I of the unit interval $[0, 1)$ we have

$$\lim_{N\to\infty} \frac{\#\{n \mid n \le N, \langle n\alpha\rangle \in I\}}{N} = |I|, \tag{1}$$

where $\langle x\rangle$ denotes the fractional part of x, that is $\langle x\rangle = x - [x]$, and $|I|$ is the length of the interval I. In fact, Weyl went on to prove, in [Weyl, 1916], that the sequence $\alpha, 2^2\alpha, 3^2\alpha \dots$ is uniformly distributed mod 1. A bit less than twenty years later Vinogradov [cf. Ellison & Ellison, 1985] proved, as a byproduct of his solution of the 'odd' Goldbach conjecture[1], that the sequence $(p_n\alpha)$, where p_n denotes the n-th prime number, is uniformly distributed mod 1. On the other hand, it is easy to see that for some irrational α the sequence $(2^n\alpha)$ is not uniformly distributed mod 1.

Now the question arises what happens if we replace the interval I in (1) by an arbitrary Lebesgue measurable subset of $[0, 1)$. What kind of extensions do the results of Weyl and Vinogradov have in this direction? We cannot expect a word-for-word generalization of their results since the Lebesgue measure of any fixed sequence is 0, and I may even be disjoint from the sequence! In the beginning of the 30's Birkhoff [1931] and Khintchin [1933] proved the appropriate generalization of (1): for any fixed Lebesgue measurable $I \subset [0, 1)$, for almost every x we have

$$\lim_{N\to\infty} \frac{\#\{n \mid n \le N, \langle x + n\alpha\rangle \in I\}}{N} = |I|, \tag{2}$$

where now $|I|$ denotes the Lebesgue measure of I. (It is not at all clear at first sight, but the result in (2) *does* imply the one in (1).) This is an instance of the individual (or pointwise) ergodic theorem. But then it took more than fifty years to obtain similar generalizations of the other result of Weyl and the result of Vinogradov! Bourgain [1988, 1988a, 1989] developed a very powerful method with which he proved in 1987: for any fixed Lebesgue measurable $I \subset [0, 1)$, for almost every x we have

$$\lim_{N\to\infty} \frac{\#\{n \mid n \le N, \langle x + n^2\alpha\rangle \in I\}}{N} = |I|, \tag{3}$$

and

$$\lim_{N\to\infty} \frac{\#\{n \mid n \le N, \langle x + p_n\alpha\rangle \in I\}}{N} = |I| \tag{4}$$

[1] That every large enough odd number is a sum of three prime numbers.

(recall that p_n is the n-th prime). Bourgain's method is a wonderful blend of (analytic) number theory, Fourier analysis and ergodic theory: he uses estimates on the Fourier transforms (or exponential sums)

$$\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i n^2 t}$$

and

$$\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i p_n t}$$

respectively, as they were obtained even by Weyl and Vinogradov.

In this paper we will examine the following general question: for what sequences of integers (a_n) do we have that for any fixed Lebesgue measurable $I \subset [0,1)$, for almost every x

$$\lim_{N\to\infty} \frac{\#\{n \mid n \le N, \langle x + a_n\alpha\rangle \in I\}}{N} = |I|? \tag{5}$$

While our primary goal is to introduce the reader to Bourgain's method that led to the results in (3) and (4), we will also discuss other methods and results of related interest, in particular the "early" results of Krengel, Bellow and others. In fact, the problem formulated in (5) is only the starting point of our investigations. To give an idea of some of the questions we shall examine, let us give the analytic reformulation of the results above.

Let f denote the characteristic function of the interval $I \subset [0,1)$. Then (1) can be rewritten as

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(\langle n\alpha\rangle) = \int_0^1 f(x)dx, \tag{1'}$$

and an approximation argument shows that in the above we can take f to be any Riemann integrable function defined on $[0,1)$.

Similarly, if f denotes the indicator function of the Lebesgue measurable set $I \subset [0,1)$, then (2) can be written as

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(\langle x + n\alpha\rangle) = \int_0^1 f(x)dx. \tag{2'}$$

Now Khintchin showed that in the above we can take f to be any Lebesgue integrable funcion. How about a similar generalization of (3) and (4)? That is, writing

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(\langle x + n^2\alpha\rangle) = \int_0^1 f(x)dx, \tag{3'}$$

and

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(\langle x + p_n\alpha\rangle) = \int_0^1 f(x)dx, \tag{4'}$$

can we take f to be any Lebesgue integrable function? The answer is not known! All we know is that (3') and (4') hold for every L^p for $p > 1$. So we arrive at the following question: let (a_n) be a sequence of integers (or real numbers). For what values of p, $1 \le p \le \infty$, do we have that for each $f \in L^p$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(\langle x + a_n\alpha\rangle) = \int_0^1 f(x)dx \tag{6}$$

for almost every x? If for some p for every $f \in L^p$ we have (6) a.e., then we call the sequence (a_n) a *universally good averaging sequence* (for L^p), because there is no restriction on the irrational α.

We have already mentioned that for some irrational α the sequence $(2^n\alpha)$ is not uniformly distributed mod 1. But we also have a result of Weyl which says that for almost every α the sequence $(2^n\alpha)$ *is* uniformly distributed mod 1. So at least there is something good to be said here. The picture changes dramatically in the ergodic-theoretical setting. It is a result of Bellow that for *every* irrational α there is a characteristic function f for which

$$\frac{1}{N} \sum_{n=1}^{N} f(\langle x + 2^n\alpha\rangle)$$

diverges for almost every x. We can say that in a sense the sequence (2^n) is a *universally bad averaging sequence.*

The methods — mostly developed by Bourgain — to solve the above "subsequence" problems helped to settle other almost everywhere convergence problems. We will mention a number of these problems, but we will develop the method in the special context of these subsequence problems, and usually we refer to other type of applications in the notes after the chapters.

Prerequisites

We do not assume that the reader is familiar with any deeper theories, but this does not mean that she/he has an easy task. The difficulty is that we use tools and results from five branches of mathematics.

Thorough knowledge of the Lebesgue integral is certainly assumed, as well as the elements of functional analysis, such as Baire's category theorem and the uniform boundedness principle. The books [Wheeden, Zygmund, 1977] or [Royden, 1988] contain all the material that is needed from measure theory, and Chapters 7 and 9 of [Royden, 1988] have all the functional analysis we need.

We will use basic facts about harmonic analysis on the classical groups $\mathbb{T}$, $\mathbb{Z}$ and $\mathbb{R}$. We will also use facts about the Hardy-Littlewood maximal function (both on the integers and on the real line), although we will give a proof on $\mathbb{Z}$, since we will use it to prove the ergodic maximal inequality. At certain points — not crucial — we will use the M. Riesz interpolation theorem. Herglotz's theorem on the representation of positive definite sequences is used. All the material that is needed from harmonic analysis can be found in the first two chapters of [Helson, 1991]. Straightforward treatments of the Hardy-Littlewood maximal function can be found in section 4.6 of [Helson, 1991] or in Chapter 9 of [Wheeden, Zygmund, 1977].

From number theory we use the basic properties of congruences. Some of the examples and exercises will refer to deep results of analytic number theory, such as the prime number theorem for arithmetic progressions, but skipping these will not seriously affect the reader's understanding of the other parts of the material. All the material needed from elementary number theory can be found in the first eleven chapters of [Weil, 1979].

The material we assume to be known from probability theory is, in addition to elementary concepts, the moment estimate of Marcinkiewicz and Zygmund. This inequality is a generalization of Khintchin's inequality for the Rademacher functions. The facts we use from probability theory can be found in the first chapter of [Durrett, 1991]. The moment estimate of Marcinkiewicz and Zygmund is in [Garsia, 1970].

Although we will give proofs of both the mean and pointwise ergodic theorems, it is desirable that the reader have some idea of the significance of these results. The reader should know what an aperiodic and an ergodic transformation is, and should know about Rokhlin's tower construction. Strictly speaking, all the material we need from ergodic theory is the first 7 sections (the seventh is "Consequences of ergodicity") and the section "Uniform topology" from [Halmos, 1956]

Summing up: the material is accesible for a third–year (US) graduate student, but she/he may want to skip some of the examples at first reading.

How to use these notes

Our presentation is fairly concise. Although we will try to give careful explanations of the underlying ideas of each proof, we will leave the routine computations to the reader.

There are exercises throughout the text, not just at the ends of the sections. Some of the exercises are difficult, but we provide hints for most of them. The exercises form an integral part of the text, and they often contain interesting developments of the preceding results.

We hope that these notes will motivate the reader to think about some of the unsolved problems of this field (many of which are mentioned in the sequel).

Finally, it is important to note the following. Even with the assumed prerequisites we could not give a full account of Bourgain's method; we give his fundamental inequality without proof, and we refer the interested reader either to Bourgain's original paper [1989] or to Thouvenot's [1990]

Acknowledgements

We want to thank all the mathematicians whose work is discussed in the text; we have freely quoted from the work of Professors M. Akcoglu, A. Bellow, M. Boshernitzan, J. Bourgain, Y. Huang, A. del Junco, R. L. Jones, K. Reinhold-Larsson and H. White. We have benefitted from discussions with Professor M. Lacey.

We also want to thank Professors K. Petersen and I. Salama for inviting us to prepare this article as a more detailed exposition of the survey lecture given at the conference "*Ergodic Theory and Its Connections with Harmonic Analysis*" held in Alexandria, Egypt, in May of 1993.

List of Symbols

We give the page number where the symbol first occurs. We explain the meaning of the symbol or notation only if a brief explanation is possible.

$\mathbb{Z}$: the set of integers	8
$\mathbb{T}$: the torus; $\mathbb{R}/\mathbb{Z}$	8
$\mathbb{R}$: the set of real numbers	8
$\mathbb{C}$: the set of complex numbers	14
$\mathbb{Z}_+$: the set of positive integers	16
$\mathbb{R}_+$: the set of positive real numbers	21
$\mathbb{N}$: the natural numbers	107
$\log x$: natural logarithm of x	27
$[x]$: largest integer not exceeding x	5
$\langle x\rangle$: fractional part of x; $\langle x\rangle = x - [x]$	5
$\#A$: the cardinality of the set A	5
(X, Σ, m, τ): dynamical system	14
$A(t)$: $A(t) = \#\{a \mid a \in A, 1 \le a \le t\}$ for $A \subset \mathbb{Z}_+$	14
$A(t,q,h)$: $= \#\{a \mid a \in A, 1 \le a \le t, a \equiv h \mod q\}$	15
(a_n): $\subset \mathbb{Z}_+$, strictly increasing (unless we say otherwise)	6
τ: a point tranformation	14
$M_t(A,f)$: $= \frac{1}{A(t)} \sum_{\substack{1\le a\le t \\ a\in A}} f(\tau^a x) = \frac{1}{t}\sum_{n\le t} f(\tau^{a_n} x)$	14
T_τ: the operator induced by τ	32
$e(y)$: $= e^{2\pi i y}$	18
$\widehat{M_t}(\beta)$: $= \frac{1}{A(t)} \sum_{\substack{1\le a\le t \\ a\in A}} e(a\beta)$	31
δ_a: the Dirac mass at the point a	15
1_B: indicator function of the set B	54
X_q: $= \{0, 1, \dots, q-1\}$	15
τ_q: $\tau_q x = x + 1 \mod q$	15
m_q: measure on X_q; $m_q(x) = 1/q$ for $x \in X_q$	15
$\\|f\\|_p = \\|f\\|_{L^p(X)}$: $= \left(\int_X \|f\|^p\right)^{1/p}$	17
$\\|f\\|_p = \\|f\\|_{\ell^p}$: $= \left(\sum_{j\in\mathbb{Z}} \|f(j)\|^p\right)^{1/p}$	49
I_ρ: (first meaning) $I_\rho = \{\rho^k\}_{k\in\mathbb{Z}_+}$ for $\rho > 1$	17
I_ρ: (second meaning)	22
$\widehat{X} = \widehat{X}_q$: $= \{0, 1/q, \dots, (q-1)/q\}$	17
$\widehat{m} = \widehat{m}_q$: measure on $\widehat{X}_q$; $\widehat{m}_q(x) = 1$	17
$\widehat{f} = \mathcal{F}f$: (on a periodic system) Fourier-transform;	17
$\widehat{f} = \mathcal{F}f$: (for $f : \mathbb{Z} \to \mathbb{C}$) $\widehat{f}(\beta) = \sum_{j\mathbb{Z}} f(j) e(-j\beta)$	31
$\check{f} = \mathcal{F}^{-1}f$: (on a periodic system) inverse F-transform;	18
$\check{f} = \mathcal{F}^{-1}f$: (for $f : \mathbb{T} \to \mathbb{Z}$) $\mathcal{F}^{-1}f(j) = \int_{\mathbb{T}} f(\beta) e(j\beta) d\beta$	82
$\overline{d}(A)$: $= \limsup_{t\to\infty} \frac{A(t)}{t}$	21
$\underline{d}(A)$: $= \liminf_{t\to\infty} \frac{A(t)}{t}$	21

List of Exercises

CHAPTER I

GOOD AND BAD SEQUENCES IN PERIODIC SYSTEMS

The ergodic-theoretical reformulation of the results in (2')-(4') rests on the observation that for a fixed real number α the transformation τ of the unit interval $[0,1)$ defined by

$$\tau(x) = \langle x+\alpha \rangle$$

is Lebesgue measurable, and preserves the (mod 1) Lebesgue measure. To fix the more general terminology to be used in the sequel we introduce

Definition 1.1. Let (X,Σ,m) be a probability measure space, and let τ be a map of X into itself. We say τ is *Σ-measurable* if and only if for each $A \in \Sigma$ we have $\tau^{-1}A \in \Sigma$. We say τ *preserves* m if and only if for each $A \in \Sigma$ we have $m(\tau^{-1}A) = m(A)$. A *measure preserving transformation* of X is a Σ-measurable map of X into itself which preserves m. Finally, the quadruple (X,Σ,m,τ), where τ is a measure-preserving tranformation of X, is called a *dynamical system.*

Let τ be a measure-preserving transformation of the probability space (X,Σ,m), and let $A=(a_n)$ be a sequence of positive integers. For $f: X \to \mathbb{C}$ consider

$$\frac{1}{N}\sum_{n=1}^{N} f(\tau^{a_n}x)\,. \tag{1.1}$$

The purpose of these notes is to examine the almost everywhere convergence of the averages in (1.1) for various sequences A when f is a characteristic function or even when it belongs to some L^p-class. In general, the a.e. limit of the averages in (1.1) is not equal to $\int_X f\,dm$ even if τ is ergodic. We will see simple examples below. Nevertheless, for irrational rotations the a.e. limit will often be the mean of f.

It turns out that instead of the averages in (1.1) it is often more instructive (and easier) to deal with

$$M_t(A,f)(x) = M_t f(x) = \frac{1}{A(t)} \sum_{\substack{1\le a\le t\\ a\in A}} f(\tau^a x) \qquad (1.2,)$$

where $A(t)$ is the counting function of A: $A(t) = \#\{a \mid a\in A, 1\le a\le t\}$. Note that $A(t)$ is defined for every — large enough — positive real number t. If the sequence (a_n) is strictly increasing then the a.e. behavior of the averages in (1.1) and (1.2) are equivalent. In fact,

in this article, unless we say otherwise, a sequence (a_n) will mean an infinite, strictly increasing sequence of positive integers.

In this chapter we examine the special case when our measure-preserving system is periodic. This means that X consists of q symbols, say $X = \{0, 1, \ldots, q-1\}$, the measure of each $x \in X$ is $1/q$, and $\tau(x) = x + 1 \mod q$. A periodic system is particularly simple, and this chapter is of illustrative nature, but some of the results will be used later. For example, in periodic systems we will even be able to prove rate of convergence results which will have important consequences for Fourier transform estimates.

In this chapter X_q denotes the set $\{0, 1, \ldots, q-1\}$; m_q denotes the measure on X_q specified by $m_q(x) = 1/q$ for $x \in X_q$; and τ_q denotes the shift transformation $\tau(x) = x + 1 \mod q$ on X_q.

Let (X, Σ, m, τ) be a periodic system on q symbols, so $X = X_q$. Every function on X is a linear combination of characteristic functions of $h \in X$. It follows that the examination of the convergence of the $M_t f$'s reduces to the examination of the convergence of the averages $M_t \delta_h$, where δ_h is the Dirac mass at the point h, i.e. the characteristic function of $\{h\}$. It is clear that it is enough to examine the convergence of the sequences $M_t \delta_h(0)$ for each $h \in X$. In other words, we just need to examine the asymptotic behavior of $A(t, q, h)/A(t)$, where

$$A(t, q, h) = \#\{a \mid a \in A, 1 \leq a \leq t, a \equiv h \mod q\},$$

that is, $A(t, q, h)$ is the number of a's that are less than or equal to t and are in the arithmetic progression $qs + h$, $s = 0, 1, \ldots$.

1. Ergodic sequences for periodic systems

In this section we give examples for sequences with the nicest property: they are uniformly distributed in residue classes for every modulus q. Since we do not want to mix this concept with the concept of uniform distribution mod 1 discussed in the introduction we use the term "ergodic sequence":

Definition 1.2. We say that the sequence of positive integers $A = (a_n)$ is *ergodic* mod q if and only if for each h we have

$$\lim_{t \to \infty} \frac{A(t, q, h)}{A(t)} = \frac{1}{q}. \tag{1.3}$$

The sequence $A = (a_n)$ is *ergodic for periodic systems* if and only if it is ergodic mod q for every q.

Note that if A is ergodic for periodic systems, then in every periodic system we have

$$\lim_{t \to \infty} M_t f(x) = \int_X f \, dm \tag{1.4}$$

for every x.

Example 1.3. The sequence of positive integers. We take $A = \mathbb{Z}_+$, the sequence of positive integers. We clearly have the estimate, since $A(t) = [t]$,

$$|A(t, q, h) - \frac{A(t)}{q}| \leq 1,$$

which implies (1.3), hence the ergodicity of $\mathbb{Z}_+$ in periodic systems.

Example 1.4. Randomly generated sequences of positive density. Here we consider random sequences. Let (Ω, β, P) be a probability space, and let $0 < \sigma < 1$. Let $Y_1, Y_2, \ldots$ be an i.i.d. sequence of 0-1 valued random variables with $P(Y_n = 1) = \sigma$ and $P(Y_n = 0) = 1 - \sigma$. Then for each $\omega \in \Omega$ we can define a sequence of positive integers A^ω by letting $n \in A^\omega$ if and only if $Y_n(\omega) = 1$.

We want to show that for almost every ω the sequence A^ω is ergodic for periodic systems.

It is enough to show that for each fixed q, with probability 1 the sequence A^ω is ergodic mod q. In anticipating the techniques used later, we give a proof which is not the simplest one; we are going to show that with probability 1 we have

$$\lim_{t\to\infty} M_t(A^\omega, f)(x) = \int_X f dm \tag{1.5}$$

for every x. By the strong law of large numbers, since the expectation of Y_n is σ, with probability 1 we have

$$\lim_{t\to\infty} \frac{A^\omega(t)}{\sigma \cdot t} = 1.$$

(The observant reader will notice that this is also incuded in (1.14) to be proved below.) It follows that for almost every ω (1.5) is equivalent with

$$\lim_{t\to\infty} Q_t^\omega f(x) = \int_X f dm, \tag{1.6}$$

where

$$Q_t^\omega f(x) = \frac{1}{\sigma t} \sum_{\substack{1 \leq a \leq t \\ a \in A^\omega}} f(\tau^a x).$$

Let us rewrite $Q_t^\omega f$ as

$$Q_t^\omega f(x) = \frac{1}{\sigma t} \sum_{n \leq t} Y_n(\omega) f(\tau^n x). \tag{1.7}$$

Then we are given the idea to compare $Q_t^\omega f$ with its "expectation"

$$V_t f(x) = \frac{1}{\sigma t}\sum_{n\le t}\sigma f(\tau^n x)\,,$$

which is just the usual ergodic average examined in the previous section. (Here we see the advantage of using the averages of (1.2) instead of (1.3).) Since X has only finitely many elements, it is enough to prove that with probability 1

$$\lim_{t\to\infty}\|Q_t^\omega f(x) - V_t f(x)\|_{L^2(X)} = 0. \tag{1.8}$$

In fact, we shall not prove (1.8), but a weaker version of it. Namely, we shall prove it only when t runs through a subsequence of the positive integers. Our first lemma tells us what subsequence we can take.

Lemma 1.5. *Let (f_n) be a sequence of nonnegative numbers. For $\rho > 1$ denote $I_\rho = \{t \mid t = \rho^k$ for some positive integer $k\}$. Suppose that for each $\rho > 1$ the averages*

$$\frac{1}{t}\sum_{n\le t} f_n$$

converge to some finite limit as t runs through the sequence I_ρ.

Then for each $\rho > 1$ this limit is the same, say L, and we have

$$\lim_{t\to\infty}\frac{1}{t}\sum_{n\le t} f_n = L\,.$$

Proof. Let us set $F_t = \frac{1}{t}\sum_{n\le t} f_n$. Let $\rho > 1$. For an arbitrary (but large enough) t let us choose the positive ineger m so that $\rho^m \le t < \rho^{m+1}$. Since the sequence (f_n) is nonnegative, we can estimate

$$F_t = \frac{1}{t}\sum_{n\le t} f_n \le \frac{1}{\rho^m}\sum_{n\le\rho^{m+1}} f_n = \rho\cdot F_{\rho^{m+1}}\,. \tag{1.9}$$

Similarly, we get

$$\frac{1}{\rho}\cdot F_{\rho^m} \le F_t\,. \tag{1.10}$$

Let us denote

$$L_\rho = \lim_{\substack{s\to\infty\\ s\in I_\rho}} F_s\,.$$

Let us choose a sequence (ρ_k) so that $\rho_k > 1$, $I_{\rho_k}\subset I_{\rho_{k+1}}$ and $\rho_k \to 1$. Clearly, $L_{\rho_k} = L_{\rho_l}$ for each k,l, so set $L = L_{\rho_k}$. Using the estimates in (1.9) and (1.10) with a $\rho\in(\rho_k)$ we obtain

$$\frac{1}{\rho}\cdot L \le \liminf_{t\to\infty} F_t \le \limsup_{t\to\infty} F_t \le \rho\cdot L\,.$$

Since $\rho\in(\rho_k)$ can be taken as close to 1 as we wish, our proof is complete. $\square$

Exercise 1.6. Show, using the lemma above, that (1.8) follows from the following: For each $\rho > 1$ there is $\Omega_\rho \subset \Omega$ with $P(\Omega_\rho) = 1$, so that for each $\omega \in \Omega_\rho$

$$\lim_{\substack{t\to\infty \\ t\in I_\rho}} \|Q_t^\omega f(x) - V_t f(x)\|_{L^2(X)} = 0\,, \tag{1.11}$$

where I_ρ is defined in the lemma. (Hint: use the lemma for the negative and positive parts of f separately)

The proof of (1.11) gives us the first opportunity to use Fourier-analysis.

For $f : X_q \to \mathbb{C}$ we define its Fourier transform $\widehat{f}$ as follows. The dual group of X_q is the set $\widehat{X}_q = \{0, 1/q, \ldots, (q-1)/q\}$ with addition mod 1. We define $\widehat{f} : \widehat{X}_q \to \mathbb{C}$ by

$$\widehat{f}(b/q) = \int_{X_q} f(x)e(-xb/q)dm = \frac{1}{q}\sum_{x=0}^{q-1} f(x)e(-xb/q)\,,$$

where we set $e(y) = e^{2\pi i y}$. Often we will use $\mathcal{F}f$ to denote $\widehat{f}$.

The inverse Fourier transform of $f : \widehat{X}_q \to \mathbb{C}$ is defined by the formula

$$\mathcal{F}^{-1}f(x) = \check{f}(x) = \int_{\widehat{X}_q} f(b/q)e(xb/q)d\widehat{m} = \sum_{b=0}^{q-1} f(b/q)e(xb/q)\,.$$

Note that the $\widehat{m}$-measure of each point in $\widehat{X}_q$ is 1. Of course this is so we can write Parseval's relation in the following form

$$\int_{X_q} |f|^2 dm = \int_{\widehat{X}_q} |\widehat{f}|^2 d\widehat{m}\,.$$

Before we rewrite (1.11) using Fourier transforms we observe that

$$\mathcal{F}\left(Q_t^\omega f\right)(b/q) = \mathcal{F}Q_t^\omega(-b/q)\cdot\mathcal{F}f(b/q) \tag{1.12}$$

and

$$\mathcal{F}\left(V_t f\right)(b/q) = \mathcal{F}\left(V_t\right)(-b/q)\cdot\mathcal{F}f(b/q)\,, \tag{1.13}$$

where

$$\mathcal{F}\left(Q_t^\omega\right)(-b/q) = \frac{1}{\sigma t}\sum_{n\le t} Y_n(\omega)e(nb/q)$$

and

$$\mathcal{F}\left(V_t\right)(-b/q) = \frac{1}{\sigma t}\sum_{n\le t} \sigma\cdot e(nb/q)\,.$$

The idea behind the formulas in (1.12) and (1.13) is, of course, that $Q_t^\omega f$ and $V_t f$ can be regarded as convolutions, and the Fourier transform of a convolution is the product of the Fourier transforms. Now using Parseval's formula and (1.12), (1.13) we can estimate as

$$\|Q_t^\omega f(x) - V_t f(x)\|_{L^2(X)} = \|\mathcal{F}\left(Q_t^\omega - V_t\right)(-b/q)\mathcal{F}f(b/q)\|_{L^2(\widehat{X})} \leq$$
$$\leq \sup_{b/q\in\widehat{X}} \left|\frac{1}{\sigma t}\sum_{n\leq t}(Y_n(\omega)-\sigma)\,e(nb/q)\right| \cdot \|\widehat{f}\|_{L^2(\widehat{X})}\,.$$

Since there are only finitely many b/q's, we see we just need to prove that for each $b/q \in \widehat{X}$, for a.e. ω we have

$$\lim_{\substack{t\to\infty\\ t\in I_\rho}} \left|\frac{1}{\sigma t}\sum_{n\leq t}(Y_n(\omega)-\sigma)\,e(nb/q)\right| = 0\,. \tag{1.14}$$

Let us introduce the random variable $Z_n = (Y_n - \sigma)e(nb/q)$. We will get (1.14) if we prove

$$\int_\Omega \left(\sum_{t\in I_\rho}\left|\frac{1}{\sigma t}\sum_{n\leq t} Z_n(\omega)\right|^2\right) dP(\omega) =$$
$$= \sum_{t\in I_\rho}\int_\Omega \left|\frac{1}{\sigma t}\sum_{n\leq t} Z_n(\omega)\right|^2 dP(\omega) < \infty\,. \tag{1.15}$$

We observe, using independence, that the Z_n's are orthogonal: $\int_\Omega Z_n \cdot \overline{Z}_j dP = 0$ if $j \neq n$. It follows that

$$\int_\Omega \left|\frac{1}{\sigma t}\sum_{n\leq t} Z_n(\omega)\right|^2 dP(\omega) = \frac{1}{\sigma^2 t^2}\int_\Omega \sum_{n\leq t}|Z_n(\omega)|^2 dP(\omega) \leq \frac{2}{\sigma^2 t}\,,$$

which implies (1.15).

Let us summarize the main ideas in the above proof. First of all we observe that it is enough to prove the convergence of the averages $M_t(A,f) = M_t f$ when t runs through a sparse sequence. Then we compare the $M_t f$'s to another sequence of averages $V_t f$ for which we previously established convergence. This comparison is done by showing that the L^2-norm of the difference $M_t f - V_t f$ is small. This is achieved — and this is the main step of the proof — by estimating the L^∞-norm of the difference of the Fourier transforms $\mathcal{F}M_t - \mathcal{F}V_t$. In case of a periodic system, since the space X_q has only finitely many points, it was a simple matter to estimate this L^∞-norm; namely, we just needed an estimate at each point $b/q \in \widehat{X}_q$.

Exercises for Example 1.4

Ex.1. Let $A = (a_n)$ be a strictly increasing sequence of positive integers.

a) Show that A is ergodic mod q if and only if for each $b \not\equiv 0 \mod q$

$$\lim_{t\to\infty} \frac{1}{A(t)} \sum_{\substack{a\le t \\ a\in A}} e(ab/q) = 0\,.$$

b) Using part a), give another proof that the random sequence A^ω discussed in Example 1.4 is ergodic mod q for every q. (This proof is the same, in principle, as the one given in the text, but the viewpoint is different; it is along the lines of Weyl's proof of his equidistribution theorem.)

Ex.2. Give yet another proof that the random sequence A^ω of Example 1.4 is ergodic mod q by considering the product space $\Omega \times X_q$ with the product measure $P \times m$ etc. (Hint: use the techniques used to establish (1.14).)

Ex.3. **a)** Show that for every $1 \le p < \infty$

$$\lim_{t\to\infty} \|Q_t^\omega f(x) - V_t f(x)\|_{L^p(X_q)} = 0.$$

(We use the notation of the text.)

b) Does the conclusion of part a) hold if $p = \infty$?

Ex.4. Let (Ω, β, P) be a probability space, and let $0 \le \sigma \le 1$. Let $Y_1, Y_2 \ldots$ be an i.i.d. sequence of $((-1)$-$1)$-valued random variables wih $P(Y_n = 1) = \sigma$ and $P(Y_n = -1) = 1 - \sigma$. Denote $a_n(\omega) = \sum_{k\le n} Y_k(\omega)$. Show that with probability 1, for $f : X_q \to \mathbb{C}$ we have

$$\lim_{t\to\infty} \frac{1}{t} \sum_{n\le t} f(\tau^{a_n(\omega)}x) = \int_{X_q} f dm_q$$

for every $x \in X_q$. (Note that the random sequence $(a_n(\omega))$ is not an increasing sequence of integers anymore even if it has a "positive drift", that is when $\sigma > 1/2$.)

Ex.5. Let $Y_1, Y_2, \ldots$ be a sequence of mean 0, orthogonal and uniformly bounded random variables on the probability space (Ω, β, P). Then $1/t \sum_{n\le t} Y_n(\omega) \to 0$ almost surely.

Example 1.7. Randomly generated sequences of 0 density. The random sequence A^ω of Example 1.4 had positive density. In this section we examine random sequences of 0 density. In fact, a random sequence of 0 density has the additional property that even its Banach-density is 0.

Definition 1.8. Let $A = (a_n)$ be a strictly increasing sequence of positive integers.

The *upper density*, the *lower density*, and *density* of A are defined respectively as

$$\overline{d}(A) = \limsup_{t\to\infty} \frac{A(t)}{t},$$

$$\underline{d}(A) = \liminf_{t\to\infty} \frac{A(t)}{t},$$

$$d(A) = \lim_{t\to\infty} \frac{A(t)}{t}.$$

The *upper Banach density*, the *lower Banach density*, and *Banach density* of A are defined respectively as

$$\overline{Bd}(A) = \limsup_{|I|\to\infty} \frac{\#\{a \mid a \in A \cap I\}}{|I|},$$

$$\underline{Bd}(A) = \liminf_{|I|\to\infty} \frac{\#\{a \mid a \in A \cap I\}}{|I|},$$

$$Bd(A) = \lim_{|I|\to\infty} \frac{\#\{a \mid a \in A \cap I\}}{|I|},$$

where I denotes (finite) subintervals of $\mathbb{R}_+$.

The random sequence A^ω will have 0 Banach density if we put the integer n into A^ω with probability σ_n, where $\sigma_n \to 0$ as $n \to \infty$.

Let (σ_n) be a sequence of positive numbers satisfying the following properties:

(i) The sequence (σ_n) is decreasing;
(ii) $\lim_{t\to\infty} \sum_{n\le t} \sigma_n = \infty$.

Property (i) will guarantee the ergodicity of the random sequence A^ω; property (ii) guarantees that A^ω has infinitely many elements.

Let (Y_n) be a sequence of independent, (0-1)-valued random variables on the probability space (Ω, β, P) so that $P(Y_n = 1) = \sigma_n$ and $P(Y_n = 0) = 1 - \sigma_n$. In other words, the expectation $E(Y_n)$ of Y_n is σ_n.

We shall show that with probability 1 the random sequence A^ω, defined by $n \in A^\omega$ if and only if $Y_n(\omega) = 1$, is ergodic for periodic systems.

So we will show that with probability 1, if $f : X_q \to \mathbb{C}$ then (1.5) holds. This time we will compare $M_t f$ with

$$V_t f(x) = \frac{1}{\sum_{n\le t} \sigma_n} \sum_{n\le t} \sigma_n f(\tau^n x).$$

The fact that $\lim_{t\to\infty} V_t f(x) = \int_X f dm$ follows from the following lemma.

Lemma 1.9. *Suppose the sequence (σ_n) of positive numbers satisfies properties (i) and (ii) above. Suppose that for some sequence (f_n) of complex numbers we have*

$$\lim_{t\to\infty}\frac{1}{t}\sum_{n\le t} f_n = \alpha\,.$$

Then we have

$$\lim_{t\to\infty}\frac{1}{\sum_{n\le t}\sigma_n}\sum_{n\le t}\sigma_n f_n = \alpha\,.$$

Proof. It is an exercise in summation by parts.

The key step in Example 1.4 was to establish (1.14), and this time we will prove a similar result, but we need to take a sparser sequence than I_ρ. Indeed, for $\rho > 1$ we take I_ρ to consist of the numbers $t_k, k = 1, 2, \ldots$ defined by $t_k = \min\{t \mid \rho^k \le \sum_{n\le t}\sigma_n < \rho^{k+1}\}$. Similarly to Example 1.4 (using an appropriate generalization of Lemma 1.5), we need to prove

$$\lim_{\substack{t\to\infty\\ t\in I_\rho}}\left|\frac{1}{\sum_{n\le t}\sigma_n}\sum_{n\le t}(Y_n(\omega)-\sigma_n)\,e(nb/q)\right| = 0\,. \tag{1.16}$$

We will get (1.16) if we prove

$$\sum_{t\in I_\rho}\int_\Omega\left|\frac{1}{\sum_{n\le t}\sigma_n}\sum_{n\le t}(Y_n(\omega)-\sigma_n)e(nb/q)\right|^2 dP(\omega) < \infty\,. \tag{1.17}$$

By the orthogonality of the random variables $(Y_n-\sigma_n)e(nb/q)$, we have

$$\int_\Omega\left|\frac{1}{\sum_{n\le t}\sigma_n}\sum_{n\le t}(Y_n(\omega)-\sigma_n)e(nb/q)\right|^2 dP(\omega) =$$

$$= \frac{1}{(\sum_{n\le t}\sigma_n)^2}\int_\Omega\sum_{n\le t}(Y_n-\sigma_n)^2 dP(\omega)\,. \tag{1.18}$$

By the definition of I_ρ, the inequality in (1.17) will follow from (1.18) and the estimate

$$\int_\Omega\sum_{n\le t}(Y_n-\sigma_n)^2 dP(\omega) \le \sum_{n\le t}\sigma_n\,. \tag{1.19}$$

But this follows since

$$\sum_{n\le t}\int_\Omega(Y_n-\sigma_n)^2 dP(\omega) = \sum_{n\le t}\left(\int_\Omega Y_n^2 dP(\omega) - \sigma_n^2\right) \le$$

$$\le \sum_{n\le t}\int_\Omega Y_n dP(\omega) = \sum_{n\le t}\sigma_n\,.$$

Exercise 1.10. Suppose that the sequence (σ_n) satisfies property (ii) but not necessarily property (i).

a) Give a necessary and sufficient condition on (σ_n) so that with probability 1 the averages $M_t(A^\omega, f)(x)$ converge for each $f : X_q \to \mathbb{C}$ and $x \in X_q$.

b) Give an example for (σ_n) such that with probability 1 the sequence A^ω is *not* ergodic, but the averages $M_t(A^\omega, f)(x)$ converge for each $f : X_q \to \mathbb{C}$ and $x \in X_q$.

2. Sequences that are good in residue classes

In this section we examine sequences that are not ergodic for periodic systems but $\lim_{t\to\infty} M_t(A, f)(x)$ exists.

Definition 1.11. We say that the sequence of positive integers $A = (a_n)$ is *good* mod q if and only if for each h the limit

$$\Lambda(h) = \Lambda(q, h) = \lim_{t\to\infty} \frac{A(t, q, h)}{A(t)} \tag{1.20}$$

exists. The sequence $A = (a_n)$ is *good for periodic systems* if and only if it is good mod q for every q.

Note that if A is good for periodic systems, then in every periodic system the limit

$$\lim_{t\to\infty} M_t f(x) \tag{1.21}$$

exists for every $x \in X_q$.

If the sequence A is good mod q, then for each $b/q \in \widehat{X}_q$ the limit

$$\widehat{\Lambda}(b/q) = \lim_{t\to\infty} \frac{1}{A(t)} \sum_{\substack{a\le t \\ a\in A}} e(ab/q) \tag{1.22}$$

exists. We have seen in Exercise 1 for Example 1.4 that A is ergodic mod q if and only if $\widehat{\Lambda}(b/q) = 0$ for each $b \not\equiv 0 \mod q$. Therefore, for a nonergodic sequence we will find a b for which $\widehat{\Lambda}(b/q) \neq 0$. Now, we have the trivial estimate $|\widehat{\Lambda}(b/q)| \le 1$ for every b. In the examples below we shall see good sequences that are not ergodic mod q; nevertheless,

$$\sup_{\substack{b/q\in\widehat{X}_q \\ \gcd(b,q)=1}} |\widehat{\Lambda}(b/q)| \to 0$$

as $q \to \infty$. So at least in the asymptotic behavior of the Fourier transform we can detect some kind of uniform distribution in residue classes.

Example 1.12. The sequence of squares; $|\widehat{\Lambda}(b/q)| \leq C/\sqrt{q}$. Here we examine the sequence of squares $A = \{n^2 \mid n \in \mathbb{Z}_+\}$. This sequence is certainly not ergodic; for example, we have $A(t,3,2) = 0$. To estimate $A(t,q,h)$, let $\theta(h) = \theta(q,h) = \#\{j \mid 0 \leq j < q, j^2 \equiv h \mod q\}$. Clearly, for any fixed j with $0 \leq j < q$ we have

$$\left| \#\{n \mid 1 \leq n \leq A(t), n \equiv j \mod q\} - \frac{A(t)}{q} \right| \leq 1 \,,$$

hence

$$\left| A(t,q,h) - \theta(h)\frac{A(t)}{q} \right| \leq \theta(h) \,. \tag{1.22}$$

It follows that A is good for periodic systems, and $\Lambda(q,h) = \theta(h)/q$.

Exercise 1.13. Let q be an odd prime number, and suppose $\gcd(q,h) = 1$. Then $\theta(q,h)$ is either 0 or 2. (Hint: the fact that $j^2 \equiv h \mod q$ has at most 2 solutions follows if we notice that the residue classes mod q form a field.)

Let us see how the lack of ergodicity is reflected in the asymptotic behavior of the Fourier transform

$$\mathcal{F}M_t(-b/q) = \frac{1}{\sqrt{t}} \sum_{n^2 \leq t} e(n^2 b/q) \,.$$

Proposition 1.14. *Suppose* $\gcd(b,q) = 1$. *Then we have the following equalities*

$$|\widehat{\Lambda}(b/q)| = \lim_{t \to \infty} \left| \frac{1}{\sqrt{t}} \sum_{n^2 \leq t} e(n^2 b/q) \right| = \begin{cases} 1/\sqrt{q} & \text{if } q \equiv \pm 1 \mod 4 \\ \sqrt{2/q} & \text{if } q \equiv 0 \mod 4 \\ 0 & \text{if } q \equiv 2 \mod 4 \,. \end{cases}$$

This proposition shows two things. First, we now see in terms of the Fourier transform that the sequence of squares is not ergodic for any $q > 1$. Second, we have some kind of uniformity of the distribution of the squares in residue classes; we have

$$\sup_{\substack{b/q \in \widehat{X}_q \\ \gcd(b,q)=1}} |\widehat{\Lambda}(b/q)| \to 0 \,.$$

Proof of Proposition 1.14. First of all, we have the following explicit formulas for $\widehat{\Lambda}(b/q)$:

$$\widehat{\Lambda}(b/q) = \frac{1}{q} \sum_{n<q} e(n^2 b/q) = \frac{1}{q} \sum_{h<q} \theta(h) e(hb/q) \,.$$

Because of periodicity by q, for any integer m we have

$$\widehat{\Lambda}(b/q) = \frac{1}{q}\sum_{n<q} e((n+m)^2 b/q)\,.$$

Using this, we can write

$$\begin{aligned}
|\widehat{\Lambda}(b/q)|^2 = \overline{\widehat{\Lambda}(b/q)} \cdot \widehat{\Lambda}(b/q) &= \frac{1}{q}\sum_{m<q} e(-m^2 b/q)\frac{1}{q}\sum_{n<q} e(n^2 b/q) = \\
&= \frac{1}{q}\sum_{m<q} e(-m^2 b/q)\frac{1}{q}\sum_{n<q} e\left((n+m)^2 b/q\right) = \\
&= \frac{1}{q}\sum_{m<q}\frac{1}{q}\sum_{n<q} e\left((n^2+2nm)b/q\right) = \\
&= \frac{1}{q}\sum_{n<q} e(n^2 b/q)\frac{1}{q}\sum_{m<q} e(2nmb/q)\,.
\end{aligned}$$

Now, $\frac{1}{q}\sum_{m<q} e(2nmb/q)$ is nonzero — and then it is equal to 1 — if and only if q divides $2n$, that is when either $n = 0$ or (in case q is even) $n = q/2$. It follows that for odd q we have

$$|\widehat{\Lambda}(b/q)|^2 = \frac{1}{q}\,,$$

and for even q we have

$$|\widehat{\Lambda}(b/q)|^2 = \frac{1}{q}\left(1 + e(qb/4)\right) = \begin{cases} \frac{2}{q}, & \text{if } 4 \text{ divides } q \\ 08 & \text{if } q \equiv 2 \mod 4\,. \end{cases}$$

□

Exercises for Example 1.12

Ex.1. Let $A = (a_n)$ be a strictly increasing sequence of positive integers. Show that A is good mod q if and only if for each $b/q \in \widehat{X}_q$ the limit

$$\lim_{t\to\infty} \frac{1}{A(t)} \sum_{\substack{a\le t \\ a\in A}} e(ab/q)$$

exists.

Ex.2. Let k be a positive integer, and let $A = (n^k)$ be the sequence of k-th powers.

a) Show that the sequence A is good for periodic systems by showing that $\Lambda(q,h) = \theta_k(q,h)/q$, where $\theta_k(q,h)$ is the number of solutions of the congruence $x^k \equiv h \mod q$.

b) Show that

$$\widehat{\Lambda}(b/q) = \frac{1}{q}\sum_{n<q} e(n^k b/q) = \frac{1}{q}\sum_{h<q} \theta_k(h) e(hb/q)\,.$$

Ex.3. Let k be a positive integer, and let $A = (n^k)$ be the sequence of k-th powers. Let q be a prime number and let b be relatively prime to q.

a) Show that for any fixed n the congruence $m^k \equiv n \mod q$ has at most k solutions.

b) Show that

$$\sum_{n<q} |\widehat{\Lambda}(n/q)|^2 \leq k^2 .$$

(Hint: Use Parseval's identity for the function $f : X_q \to \mathbb{C}$ defined by $f(x) = \theta_k(x, q)$.)

c) Show that

$$|\widehat{\Lambda}(b/q)|^2 = \frac{1}{q} \sum_{n<q} |\widehat{\Lambda}(n^k b/q)|^2.$$

(Hint: If $\gcd(n, q) = 1$, then $\widehat{\Lambda}(b/q) = \widehat{\Lambda}(n^k b/q)$.)

d) Show that

$$|\widehat{\Lambda}(b/q)| \leq \frac{k^{3/2}}{\sqrt{q}} .$$

Ex.4. Let k, A, q, b be as in the previous exercise. Let $d = \gcd(k, q-1)$. Show that

$$|\widehat{\Lambda}(b/q)| \leq \frac{d}{\sqrt{q}} .$$

(Hint: Be more careful with the estimations in the previous exercise.)

We remark that the best constant in the above inequality is $d - 1$.

Ex.5. Let (σ_n) be a sequence of positive numbers satisfying the properties:

(i) the sequence (σ_n) is decreasing;
(ii) $\lim_{t\to\infty} \sum_{n\leq t} \sigma_n = \infty$.

Let (Y_n) be a sequence of independent, (0-1)-valued random variables on the probability space (Ω, β, P) so that $P(Y_n = 1) = \sigma_n$ and $P(Y_n = 0) = 1-\sigma_n$.

Let k be a fixed positive number. Show that with probability 1 the random sequence A^ω, defined by $n^k \in A^\omega$ if and only if $Y_n(\omega) = 1$, is good for periodic systems.

Ex.6. Let (σ_n) be a sequence of positive numbers satisfying the properties

(i) $\lim_{n\to\infty} \sigma_n = 0$;
(ii) $\lim_{t\to\infty} \sum_{n\leq t} \sigma_n = \infty$.

Let (Y_n) be a sequence of independent, (0-1)-valued random variables on the probability space (Ω, β, P) so that $P(Y_n = 1) = \sigma_n$ and $P(Y_n = 0) = 1-\sigma_n$.

Show that with probability 1 the random sequence A^ω, defined by $n \in A^\omega$ if and only if $Y_n(\omega) = 1$, is of 0 Banach density.

Example 1.15. The sequence of primes. Let A be the sequence of prime numbers. To describe the distribution of the primes in residue classes we introduce the function

$$li(t) = \sum_{2 \le n \le t} \frac{1}{\log n}.$$

Note that $\frac{li(t)}{t/\log t} \to 1$. The prime number theorem tells us that $\frac{A(t)}{li(t)} \to 1$. If q and h are not relatively prime, then the arithmetic progression $qn + h$, $n = 1, 2, \ldots$, contains at most 1 prime number. On the other hand, the primes are uniformly distributed among the reduced residue classes. This is the content of Siegel-Walfis's theorem [Ellison & Ellison, 1985] on the distribution of primes in arithmetic progressions. Before we state the theorem, we recall the definition of Euler's "ϕ-function": $\phi(q) = \#\{h \mid 0 \le h < q, \gcd(h,q) = 1\}$.

Lemma 1.16 (The prime number theorem for arithmetic progressions). *Let* $\gcd(q,h) = 1$. *Then we have*

$$\lim_{t \to \infty} \frac{A(t,q,h)}{li(t)/\phi(q)} = 1.$$

An immediate consequence of this lemma is that the sequence of primes is good for periodic systems and $\Lambda(q,h) = 1/\phi(q)$ *if* $\gcd(q,h) = 1$ *and* $\Lambda(q,h) = 0$ *otherwise.*

Exercises for Example 1.15

Ex.1. Suppose $\gcd(b,q) = 1$. Then we have

$$\widehat{\Lambda}(b/q) = \frac{\mu(q)}{\phi(q)},$$

where μ denotes the Möbius function:

$$\mu(q) = \begin{cases} 1, & \text{if } q = 1 \\ (-1)^k, & \text{if } q \text{ is squarefree, and } k \text{ is the number of} \\ & \quad \text{prime-divisors of } q \\ 0, & \text{otherwise}. \end{cases}$$

(Hint: use the inversion formula of Möbius.)

Ex.2.

a) Show that there is a positive constant c so that for every q

$$\phi(q) > c \cdot \frac{q}{\log q}.$$

b) Show that there is a positive constant c so that for every q

$$\phi(q) > c \cdot \frac{q}{\log \log q}.$$

Ex.3. Let (σ_n) be a sequence of positive numbers satisfying the properties

(i) the sequence (σ_n) is decreasing;
(ii) $\lim_{t\to\infty} \sum_{n\le t} \sigma_n = \infty$.

Let (Y_n) be a sequence of independent, (0-1)-valued random variables on the probability space (Ω, β, P) so that $P(Y_n = 1) = \sigma_n$ and $P(Y_n = 0) = 1-\sigma_n$.

Let p_n denote the n-th prime number. Show that with probability 1 the random sequence A^ω, defined by $p_n \in A^\omega$ if and only if $Y_n(\omega) = 1$, is good for periodic systems. (Compare this with Exercise 5 for Example 1.12. Is there a common generalization?!)

3. Sequences that are bad for periodic systems

In this section we briefly examine sequences $A = (a_n)$ that are not good for periodic systems. We may distinguish several degrees of bad behavior. It is possible that A is not good mod q for only some values of q. But maybe A is not good mod q for *every* q. If we know that $M_t(A, f)(x)$ does not converge, it may mean only that we do not have convergence only for some $x \in X_q$, but maybe there is no convergence for any $x \in X_q$. Since $0 \le A(t,q,h)/A(t) \le 1$, the worst possible scenario is when

$$\limsup_{t\to\infty} \frac{A(t,q,h)}{A(t)} = 1 \tag{1.23}$$

and

$$\liminf_{t\to\infty} \frac{A(t,q,h)}{A(t)} = 0 \tag{1.24}$$

for every $h \in X_q$.

Definition 1.17. We say that the sequence of positive integers $A = (a_n)$ is *bad* mod q if and only if for each $h \in X_q$ the limit

$$\lim_{t\to\infty} \frac{A(t,q,h)}{A(t)}$$

does not exist. The sequence $A = (a_n)$ is *bad for periodic systems* if and only if it is bad mod q for every q.

We say that the sequence of positive integers $A = (a_n)$ is *sweeping out* mod q if and only if for each $h \in X_q$ the equalities in (1.23) and (1.24) hold. The sequence $A = (a_n)$ is *sweeping out for periodic systems* if and only if it is bad mod q for every q.

Note that A is sweeping out mod q if and only if for every $x \in X_q$ we have

$$\limsup_{t\to\infty} M_t(A, \delta_0)(x) = 1, \qquad \liminf_{t\to\infty} M_t(A, \delta_0)(x) = 0\,. \tag{1.25}$$

Example 1.18. A sweeping out sequence.

Here we will show that there is a strictly increasing sequence of positive integers which is sweeping out for periodic systems.

The idea of the construction is to make sure that for each q and h there is a $t = t_{(q,h)}$ so that most of the elements $a_1, a_2, \dots, a_t$ of A are congruent to $h \mod q$. It follows, of course, that for some t' most of the elements $a_1, a_2, \dots, a_{t'}$ of A are *not* congruent to $h \mod q$. It is more convenient in this proof to consider the averages

$$M_t f(x) = \frac{1}{t} \sum_{n \leq t} f(\tau^{a_n} x),$$

and we assume that t is an integer. For each q let us define $K_q = 1 + 2 + \cdots + q$. Note that each residue class mod q has a representative k satisfying $K_{q-1} < k \leq K_q$. Let us consider a sequence (t_k) of positive integers which satisfy the following condition:

$$\frac{t_k}{t_{k-1}} > q + 1; \quad \text{for } K_{q-1} < k \leq K_q. \tag{1.26}$$

Let us define the strictly increasing sequence $A = (a_n)$ to satisfy, for every k,

$$a_n \equiv k \mod q \quad \text{for} \quad t_{k-1} < n \leq t_k. \tag{1.27}$$

Let us show (1.25); so fix $q > 1$ and let $x \in X_q$. Then there is h with $0 < h \leq q$ such that $x \equiv -h \mod q$. For each positive integer s define the indices $k(s)$ by $k(s) = K_{s \cdot q - 1} + h$. We claim that

$$\lim_{s \to \infty} M_{t_{k(s)}} \delta_0(x) = 1. \tag{1.28}$$

To see this, first we note that by (1.27) we have for every s

$$a_n \equiv h \mod q \quad \text{for} \quad t_{k(s)-1} < n \leq t_{k(s)},$$

and hence

$$x + a_n \equiv 0 \mod q \quad \text{for} \quad t_{k(s)-1} < n \leq t_{k(s)}. \tag{1.29}$$

We also have, by the property in (1.26),

$$\frac{t_{k(s)}}{t_{k(s)-1}} > s \cdot q + 1. \tag{1.30}$$

Now we can estimate, by (1.29) and (1.30),

$$M_{t_{k(s)}} \delta_0(x) \geq \frac{1}{t_{k(s)}} \sum_{t_{k(s)-1} < n \leq t_{k(s)}} \delta_0(x + a_n) = \frac{t_{k(s)} - t_{k(s)-1}}{t_{k(s)}} > 1 - \frac{1}{s \cdot q},$$

which implies (1.28). On the other hand, we claim that

$$\lim_{s\to\infty} M_{t_{k(s)+1}}\delta_0(x) = 0\,. \tag{1.31}$$

This, together with (1.28), proves (1.25). To see (1.31), note that for every $s > 1$

$$a_n \equiv h+1 \mod q \quad \text{for} \quad t_{k(s)} < n \le t_{k(s)+1}\,,$$

and hence, if $q > 1$,

$$x + a_n \not\equiv 0 \mod q \quad \text{for} \quad t_{k(s)} < n \le t_{k(s)+1}\,. \tag{1.32}$$

This time we can estimate, by (1.32),

$$M_{t_{k(s)+1}}\delta_0(x) \le \frac{1}{t_{k(s)+1}} \sum_{1\le n\le t_{k(s)}} \delta_0(x+a_n) \le \frac{t_{k(s)}}{t_{k(s)+1}} < \frac{1}{s\cdot q}\,,$$

finishing the proof of (1.31).

Exercises for Example 1.18

Ex.1. Construct a sequence that is bad mod q for some q but which is good mod q for infinitely many q's.

Ex.2. Is it true that if a sequence is bad mod q for some q then it is bad mod q for infinitely many q?

Ex.3. Is it true that if a sequence is sweeping out mod q for some q then it is sweeping out mod q for infinitely many q?

Ex.4. Construct a sequence of positive density that is sweeping out for some q .

Ex.5. Is there a sequence of positive lower density that is bad for periodic systems?

Ex.6. Let (b_n) be a strictly increasing sequence of positive integers with the property that $b_{n+1} - b_n \to \infty$ as $n \to \infty$. Construct a sequence (a_n) that is bad for periodic systems and satisfies $b_n \le a_n \le b_{n+1}$ for every n.

Ex.7. **a)** Show that there is a lacunary sequence of integers which is good for periodic systems. (A sequence (a_n) is *lacunary* if and only if there is a constant $\rho > 1$ such that $a_{n+1}/a_n > \rho$ for every large enough n.)

b) Is there a lacunary sequence of integers which is ergodic for periodic systems? (Hint: there is ...)

CHAPTER II

GOOD SEQUENCES FOR MEAN L^2 CONVERGENCE

In this chapter we will examine the L^2-norm convergence of the averages $M_t(A, f)(x)$ for various sequences $A = (a_n)$. We recall that for a given sequence of positive integers $A = (a_n)$ we have defined

$$M_t(A, f)(x) = M_t f(x) = \frac{1}{A(t)} \sum_{\substack{1 \le a \le t \\ a \in A}} f(\tau^a x), \tag{2.1}$$

where τ is a measure-preserving transformation of the probability space (X, Σ, m), and $f : X \to \mathbb{C}$.

In the previous chapter we have seen that Fourier - transform estimates played an important role in establishing convergence in periodic systems, but all the results could have been proved without any reference to Fourier analysis. The role of Fourier analysis will be more significant in this chapter. Indeed, the only known proof of the mean convergence of the averages $M_t(A, f)$ for a random sequence or for the sequence of primes is via Fourier analysis. As we will see, the L^2-convergence of the averages $M_t(A, f)$ is *equivalent* with the convergence of the Fourier transforms

$$\widehat{M_t}(\beta) = \frac{1}{A(t)} \sum_{\substack{1 \le a \le t \\ a \in A}} e(a\beta)$$

for every real β.

1. Ergodic sequences

Definition 2.1. Let τ be an ergodic measure-preserving transformation of the probability space (X, Σ, m). We say that the sequence of positive integers $A = (a_n)$ is *ergodic* for the system (X, Σ, m, τ) if and only if for each $f : X \to \mathbb{C}$ we have

$$\lim_{t \to \infty} M_t(A, f) = \int_X f \, dm \tag{2.2}$$

in L^2-norm. The sequence $A = (a_n)$ is *ergodic* if and only if it is ergodic for every ergodic system (X, Σ, m, T).

Example 2.2. The mean ergodic theorem.

In this section we shall prove that the sequence of positive integers $A = \mathbb{Z}_+$ *is ergodic.*

We shall give two different proofs. The first is due to von Neumann [1932], and uses Fourier analysis; the second proof is due to F. Riesz [Riesz, Sz.-Nagy, 1990], and it uses geometric (Hilbert space) techniques. Although the first method of proof is more important for our purposes, the main idea of the second proof will also be used later.

Both proofs depend on the observation that the operator T_τ defined by $T_\tau f(x) = f(\tau x)$ is a contraction of the Hilbert space $L^2(X, \Sigma, m)$. Therefore the following should be proved:

Proposition 2.3. *Let* U *be any contraction of the Hilbert space* H. *Then for each* $f \in H$ *we have*

$$\lim_{t\to\infty} M_t f = Pf\,,$$

where $M_t f = (1/t)\sum_{n\le t} U^n f$ *and* P *is the projection onto the space of* U*-invariant points of* H.

This result implies the ergodicity of the sequence of the natural numbers, for if τ is ergodic then the τ - invariant functions are the constants, and hence, as can be seen easily, $Pf = \int_X f dm$.

Both in the Fourier analysis and in the geometric proof we just prove the existence of the limit. For the identification of the limit — that it is Pf — we refer the reader to Exercise 5 below.

Fourier analysis proof. We want to prove that the directed set $\{M_t f \mid t \in \mathbb{R}_+\}$ is "Cauchy", that is

$$\lim_{t,s\to\infty} \|M_t f - M_s f\|_H = 0\,. \tag{2.3}$$

Let (f, g) denote the inner product on H. Let us define the sequence of operators $(T_n \mid n = 0, \pm1, \pm2, \dots)$ by $T_n = U^n$ for $n \ge 0$ and $T_n = (U^*)^{-n}$ for $n < 0$. Then for a fixed $f \in H$ the sequence of real numbers $(x_n \mid n = 0, \pm1, \pm2, \dots)$ defined by $x_n = (T_n f, f)$ is a positive - definite sequence (cf. Section 9 of the Appendix in [Riesz, Sz.-Nagy, 1990]), hence, by Herglotz's theorem, there is a (nonnegative) Borel measure $\mu = \mu_f$ on the interval $[0, 1)$ so that for every $n \in \mathbb{Z}$

$$(T_n f, f) = \int e\,(n\beta)\, d\mu(\beta)\,. \tag{2.4}$$

Note that we have, in particular, $\int d\mu = \|f\|^2$. A consequence of (2.4) is the following fundamental inequality of von Neumann:

Exercise 2.4. Let $p(x)$ be a polynomial with complex coefficients. Then we have, for every $f \in H$,

$$\begin{aligned} \|p(\tau)f\|^2 &\leq \int |p\,(e(\beta))\,|^2 d\mu(\beta) \leq \\ &\leq \sup_\beta |p\,(e(\beta))\,|^2 \cdot \|f\|^2 . \end{aligned} \tag{2.5}$$

(Hint: Use (2.4) and induction on the degree of $p(x)$.)

By this exercise, we have

$$\|M_t f - M_s f\|^2 \leq \int |\widehat{M_t}(\beta) - \widehat{M_s}(\beta)|^2 d\mu(\beta)\,,$$

where

$$\widehat{M_t}(\beta) = \frac{1}{t} \sum_{1 \leq n \leq t} e(n\beta)\,.$$

Observing that $e(n\beta)$, $n = 1, 2, \ldots$, is a geometric progression, we get that $\widehat{M_t}(\beta) \to 0$ for $\beta \in [0,1) \setminus \{0\}$, hence $\widehat{M_t}(\beta) - \widehat{M_s}(\beta) \to 0$ for every β. But then, by the bounded convergence theorem, we get (2.3). □

Geometric Proof. Let I denote the space of U-invariant points of H, and let C be the set of coboundaries, that is, $C = \{f \mid f = g - Ug \text{ for some } g \in H\}$. It is clear that for $f \in I$ we have $M_t f \to f$, and if $f \in C$ then $M_t f \to 0$. By the uniform boundedness principle, we just need to prove that the linear span $I + C = \{f + g \mid f \in I, g \in C\}$ of the invariant points and coboundaries is dense in H. We shall prove this by showing that the orthocomplement $C^\perp$ of C contains only U-invariant points.

Suppose that h is orthogonal to all points f of the form $f = g - Ug$. We want to show that h is U-invariant. By assumption we have $(g - Ug, h) = 0$ for every $g \in H$. In particular, we have $(h - Uh, h) = 0$. But then, by the Pythagorean theorem, and using that U is a contraction, we get

$$\|h - Uh\|_H^2 = \|Uh\|_H^2 - \|h\|_H^2 \leq 0\,,$$

which implies that $h = Uh$. □

Exercises for Example 2.2

Ex.1. Let $A = (a_n)$ a "block" sequence. This means that A is a union of intervals of consecutive integers: $A = \cup I_k \cap \mathbb{Z}_+$, where (I_k) is a sequence of intervals in $\mathbb{R}_+$.

Show, using both the Fourier - analytic and the geometric method, that A is ergodic provided $|I_k| \to \infty$ as $k \to \infty$.

Ex.2. Let $F(t)$ be a real function satisfying the properties

(i) $F(t) \leq t$;
(ii) $\lim_{t\to\infty} F(t) = \infty$.

a) Show that there is a strictly increasing sequence A of positive integers which is ergodic and satisfies

$$\lim_{t\to\infty} \frac{A(t)}{F(t)} = 0 .$$

(Hint: use the result of the previous exercise.)

b) Is there is a strictly increasing sequence A of positive integers which is ergodic and satisfies

$$\lim_{t\to\infty} \frac{A(t)}{F(t)} = 1 ?$$

Ex.3. Let $A = (a_n)$ be a strictly increasing sequence of positive integers. Show that A is ergodic if and only if for each $\beta \in (0,1)$

$$\lim_{t\to\infty} \frac{1}{A(t)} \sum_{\substack{a\leq t\\ a\in A}} e(a\beta) = 0 .$$

Ex.4. a) Let $1 \leq p < \infty$. Prove that $M_t(\mathbb{Z}_+, f) \to \overline{f}$ in L^p norm, where $\overline{f}$ is τ-invariant. (The $L^p \to L^p$ operator T is defined by $Tf(x) = f(\tau x)$.)

b) Can we take $p = \infty$ in part a)?

Ex.5. Finish the proof of Proposition 2.3 by identifying $\lim_{t\to\infty} M_t f$ as Pf. With the notation of the geometric proof, this means that you have to show that, in fact, $C^\perp = I$. (Hint: In the geometric proof we showed that $C^\perp \subseteqq I$. To show that $I \subseteq C^\perp$ write $f \in I$ as $f = g + h$ with $g \in C^\perp$ and $h \in \overline{C}$. Then, since $h \in I \cap \overline{C}$, we have $h = \lim_{t\to\infty} M_t h = 0$.)

Example 2.5: Randomly generated sequences of positive density. Here we examine the random sequence considered in Example 1.4. So let (Ω, β, P) be a probability space, and let $0 < \sigma < 1$. Let $Y_1, Y_2, \ldots$ be an i.i.d. sequence of 0-1 valued random variables with $P(Y_n = 1) = \sigma$ and $P(Y_n = 0) = 1 - \sigma$. For $\omega \in \Omega$ we define the sequence of positive integers A^ω by letting $n \in A^\omega$ if and only if $Y_n(\omega) = 1$.

We want to show that for almost every ω the sequence A^ω is ergodic.

As in Example 1.4, we want to show that with probability 1 we have, for every $f \in L^2$,

$$\lim_{t\to\infty} \|Q_t^\omega f(x) - V_t f(x)\|_{L^2(X)} = 0 ,$$

where

$$Q_t^\omega f(x) = \frac{1}{\sigma t} \sum_{n\leq t} Y_n(\omega) f(\tau^n x) \tag{2.6}$$

and

$$V_t f(x) = \frac{1}{\sigma t} \sum_{n \leq t} \sigma f(\tau^n x).$$

Again, we just need to show that for each fixed $\rho > 1$, with probability 1 we have, for every $f \in L^2$,

$$\lim_{\substack{t \to \infty \\ t \in I_\rho}} \|Q_t^\omega f(x) - V_t f(x)\|_{L^2(X)} = 0, \tag{2.7}$$

where $I_\rho = \{t \mid t = \rho^k \text{ for some positive integer } k\}$. By the second inequality of Exercise 2.4 we need to prove

$$\lim_{\substack{t \to \infty \\ t \in I_\rho}} \sup_{\beta \in [0,1)} |\widehat{Q}_t^\omega(\beta) - \widehat{V}_t(\beta)| = 0, \tag{2.8}$$

where

$$\widehat{Q}_t^\omega(\beta) = \frac{1}{\sigma t} \sum_{n \leq t} Y_n(\omega) e(n\beta), \qquad \widehat{V}_t(\beta) = \frac{1}{\sigma t} \sum_{n \leq t} \sigma e(n\beta).$$

In Example 1.4 we computed second moments to establish what we wanted (cf. (1.15)). Here we need to take 4-th moments; we will prove that

$$\int_\Omega \left(\sum_{t \in I_\rho} \sup_{\beta \in [0,1)} |\widehat{Q}_t^\omega(\beta) - \widehat{V}_t(\beta)|^4 \right) dP(\omega)$$
$$= \sum_{t \in I_\rho} \int_\Omega \sup_{\beta \in [0,1)} |\widehat{Q}_t^\omega(\beta) - \widehat{V}_t(\beta)|^4 dP(\omega) < \infty. \tag{2.9}$$

The reason for this complication is that in proving (2.9) we cannot get by with estimates for a single fixed β from the dual group as we did in Example 1.4, because this time the "dual group" — the interval $[0,1)$ — has infinitely many elements. But it turns out we do not really have to get estimations for every $\beta \in [0,1)$. We will use the fact that $\widehat{Q}_t^\omega(\beta) - \widehat{V}_t(\beta)$ is a trigonometric polynomial of degree $[t]$ and hence cannot change too rapidly. As a consequence of Bernstein's theorem on the derivative of a trigonometric polynomial, it is enough to obtain estimates on $|\widehat{Q}_t^\omega(\beta) - \widehat{V}_t(\beta)|$ for roughly t well chosen $\beta's$ to get a uniform estimate for every $\beta \in [0,1)$.

We shall prove the following: there is a constant C (< 100) so that for every β and t

$$\int_\Omega |\widehat{Q}_t^\omega(\beta) - \widehat{V}_t(\beta)|^4 dP(\omega) \leq \frac{C}{t^2}. \tag{2.10}$$

Let us see how one can finish the proof, assuming (2.11). Let us denote $S_t = \{0, 1/10t, 2/10t, \ldots, (10t-1)/10t\}$. By (2.11), we have the estimate

$$\int_\Omega \sup_{\beta \in S_t} |\widehat{Q}_t^\omega(\beta) - \widehat{V}_t(\beta)|^4 dP(\omega) \leq \frac{C}{t}. \tag{2.11}$$