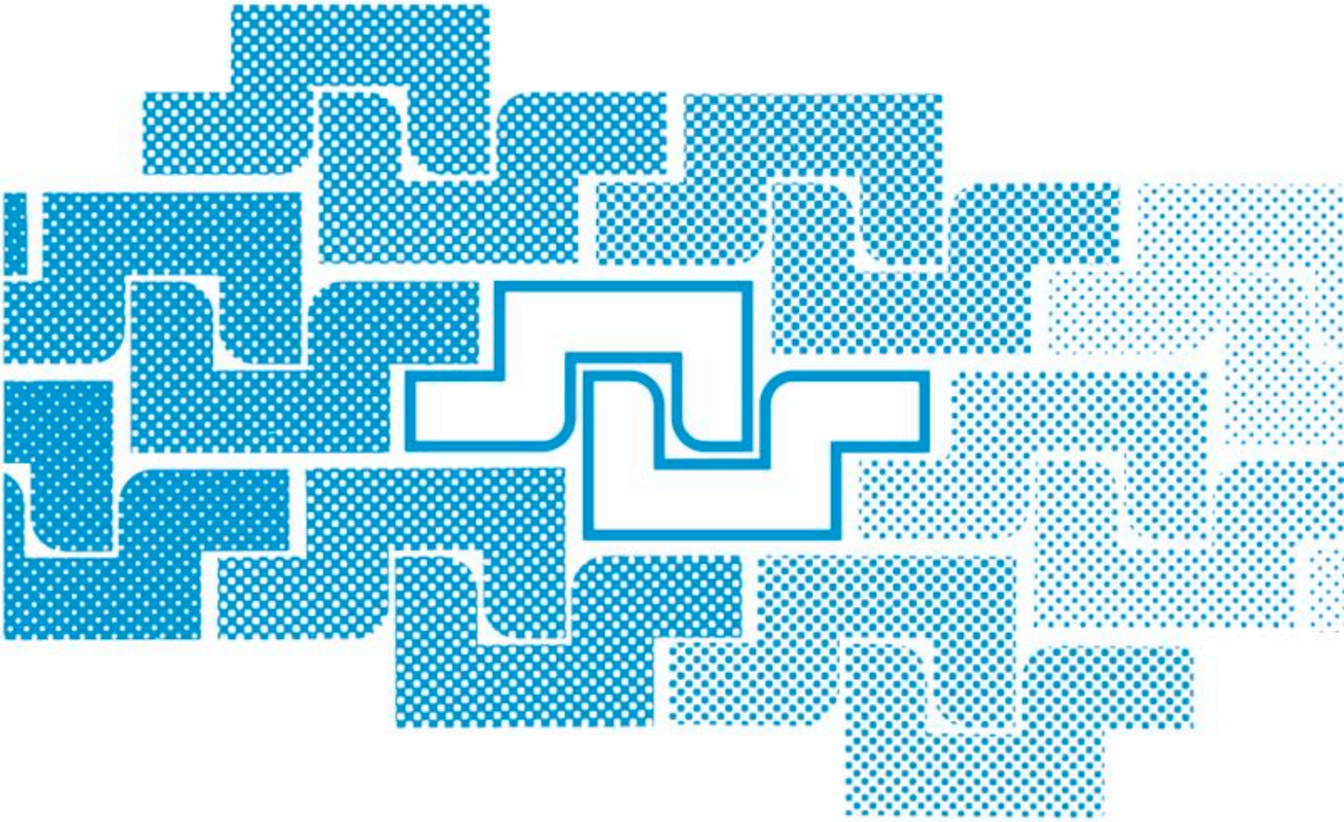


Equidistribution in Number Theory, An Introduction

Edited by

Andrew Granville and Zeév Rudnick

NATO Science Series

II. Mathematics, Physics and Chemistry – Vol. 237

Equidistribution in Number Theory, An Introduction

NATO Science Series

A Series presenting the results of scientific meetings supported under the NATO Science Programme.

The Series is published by IOS Press, Amsterdam, and Springer in conjunction with the NATO Public Diplomacy Division

Sub-Series

I. Life and Behavioural Sciences	IOS Press
II. Mathematics, Physics and Chemistry	Springer
III. Computer and Systems Science	IOS Press
IV. Earth and Environmental Sciences	Springer

The NATO Science Series continues the series of books published formerly as the NATO ASI Series.

The NATO Science Programme offers support for collaboration in civil science between scientists of countries of the Euro-Atlantic Partnership Council. The types of scientific meeting generally supported are “Advanced Study Institutes” and “Advanced Research Workshops”, and the NATO Science Series collects together the results of these meetings. The meetings are co-organized bij scientists from NATO countries and scientists from NATO’s Partner countries – countries of the CIS and Central and Eastern Europe.

Advanced Study Institutes are high-level tutorial courses offering in-depth study of latest advances in a field.
Advanced Research Workshops are expert meetings aimed at critical assessment of a field, and identification of directions for future action.

As a consequence of the restructuring of the NATO Science Programme in 1999, the NATO Science Series was re-organised to the four sub-series noted above. Please consult the following web sites for information on previous volumes published in the Series.

http://www.nato.int/science
http://www.springer.com
http://www.iospress.nl

Series II: Mathematics, Physics and Chemistry – Vol. 237

Equidistribution in Number Theory, An Introduction

edited by

Andrew Granville
University of Montréal, QC,
Canada

and

Zeév Rudnick
Tel-Aviv University,
Israel

Springer

Published in cooperation with NATO Public Diplomacy Division

Proceedings of the NATO Advanced Study Institute on
Equidistribution in Number Theory
Montréal, Canada
11–22 July 2005

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-5403-3 (PB)
ISBN-13 978-1-4020-5403-7 (PB)
ISBN-10 1-4020-5402-5 (HB)
ISBN-13 978-1-4020-5402-0 (HB)
ISBN-10 1-4020-5404-1 (e-book)
ISBN-13 978-1-4020-5404-4 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

CONTENTS

PREFACE

From July 11th to July 22nd, 2005, a NATO advanced study institute, as part of the series “Séminaire de mathématiques supérieures”, was held at the Université de Montréal, on the subject *Equidistribution in the theory of numbers*. There were about one hundred participants from sixteen countries around the world. This volume presents details of the lecture series that were given at the school.

Across the broad panorama of topics that constitute modern number theory one finds shifts of attention and focus as more is understood and better questions are formulated. Over the last decade or so we have noticed increasing interest being paid to distribution problems, whether of rational points, of zeros of zeta functions, of eigenvalues, etc. Although these problems have been motivated from very different perspectives, one finds that there is much in common, and presumably it is healthy to try to view such questions as part of a bigger subject. It is for this reason we decided to hold a school on “Equidistribution in number theory” to introduce junior researchers to these beautiful questions, and to determine whether different approaches can influence one another.

There are far more good problems than we had time for in our schedule. We thus decided to focus on topics that are clearly inter-related or do not require a lot of background to understand. Since there were two major number theory research programs taking place during the academic year 2005–2006, on *Analysis in number theory* at the Centre de recherches mathématiques in Montréal, and during the spring semester 2006 on *Rational and Integral Points on Higher-Dimensional Varieties* at the Mathematical Sciences Research Institute in Berkeley, California, we decided to help prepare junior participants by inviting some lecturers who would go on to be senior participants at those programs.

The lectures split into roughly ten topics (with lecturers):

1. The basics of uniform distribution (Granville, Rudnick).
2. Exponential Sums and cryptography (Friedlander, Granville).
3. Spectral Theory (Venkatesh).

4. Hyperbolic geometry, Ergodic theory and Ratner's theorem (Marklof, Lindenstrauss, Venkatesh).
5. Quantum equidistribution (De Bièvre, Lindenstrauss, Rudnick, Venkatesh).
6. Distribution of integers and uncertainty principles (Soundararajan, Granville).
7. Distribution of rational points on varieties (Heath-Brown, Tschinkel).
8. Distribution of special points (Rudnick, Granville, Duke, Ullmo).
9. Spacing statistics (Marklof).
10. Invited lectures on relevant subjects (Harcos, Yafaev).

The lecturers were requested to try to keep their lectures mostly self-contained (though they were allowed to require some light background reading before the meeting); in particular they were asked to avoid the use of very technical terms without giving independent motivation during their talks. In general the lecturers did so, and this was reflected by the high attendance throughout the meeting, despite the temptations of Montréal in mid-summer. We have asked the lecturers to bring over that attitude to the preparation of their contributions herein.

We would like to thank the lecturers for their superb talks, and the generosity with which they worked with the participants, as well as typed up their talks for these proceedings.

We would also like to thank the *Security through Science* programme of NATO, the Centre de recherches mathématiques, the Institut des sciences mathématiques and the Université de Montréal, for their generous and willing support of this school.

The meeting would not have been possible without the organizational skills of Diane Bélanger, for which we are very grateful. This book has benefitted from the typesetting skills of Louise Letendre and André Montpetit for which we thank them.

Andrew Granville
Zeév Rudnick

CONTRIBUTORS

Stephan De Bièvre
UFR de mathématiques
Université des Sciences et Technologies de Lille
59655 Villeneuve d'Ascq
France
Stephan.De-Bievre@math.univ-lille1.fr

Ulrich Derenthal
Mathematisches Institut
Universität Göttingen
Bunsenstraße 3-5
37073 Göttingen
Germany
ulrich.derenthal@stud.uni-goettingen.de

W. Duke
Mathematics Department
UCLA
Box 951555
Los Angeles, CA 90095-1555
USA
wdduke@ucla.edu

John B. Friedlander
Department of Computer and Mathematical Sciences
University of Toronto at Scarborough
1265 Military Trail
Toronto, ON M1C 1A4
Canada
frdlndr@utsc.utoronto.ca

Andrew Granville
Département de mathématiques et de statistique
Université de Montréal
C.P 6128, succ. Centre-ville
Montréal, QC H3C 3J7
Canada
andrew@dms.umontreal.ca

D. R. Heath-Brown
Oxford University
Mathematical Institute
24-29 St. Giles'
Oxford, OX1 3LB
UK
rhb@maths.ox.ac.uk

Elon Lindenstrauss
Department of Mathematics
Fine Hall, Washington Road
Princeton University
Princeton, NJ 08544
USA

Jens Marklof
School of Mathematics
University of Bristol
Bristol BS8 1TW
UK
j.marklof@bristol.ac.uk

Zeev Rudnick
School of Mathematical Sciences
Tel-Aviv University
Schrieber Building, Room 316
Tel-Aviv 69978
Israel
rudnick@math.tau.ac.il

K. Soundararajan
Stanford University
Mathematics, Bldg. 380
450 Serra Mall
Stanford, CA 94305-2125
ksound@math.stanford.edu

Yuri Tschinkel
Mathematisches Institut
Universität Göttingen
Bunsenstraße 3-5
37073 Göttingen
Germany
yuri@uni-math.gwdg.de

Emmanuel Ullmo
Arithmétique et Géométrie Algébrique
Université Paris-Sud
Bâtiment 425
91405 Orsay Cedex
France
Emmanuel.Ullmo@math.u-psud.fr

Akshay Venkatesh
Department of Mathematics
Courant Institute of Mathematical Science
251 Mercer Street
New York, NY 10012-1185
venkatesh@cims.nyu.edu

BIOGRAPHICAL SKETCHES OF THE LECTURERS

Stephan De Bièvre is a mathematical physicist who received his education at the Universities of Antwerp, Leuven and Rochester. After a postdoctoral stay at the University of Toronto, and several years at the Université Paris 7, he became, in 1996, Professor of Mathematics at the Université de Lille 1. He developed an interest in quantum chaos and hence in eigenfunction equidistribution a little more than a decade ago.

William Duke received his PhD from the Courant Institute and was Peter Sarnak's third student. After being at UCSD for two years he moved to Rutgers as an NSF Postdoc to work with Henryk Iwaniec. He was tenured at Rutgers and since 2000 he has been at UCLA. His research interests include the analytic theory of automorphic forms, L-functions, quadratic forms and elliptic curves. He has been especially interested in extensions and analogues of classical distribution problems from analytic number theory.

John Friedlander received his doctorate at Penn State under the direction of S. Chowla and then did postdoctoral studies as Assistant to A. Selberg at the Institute for Advanced Study. He is currently University Professor of Mathematics at the University of Toronto. His primary research interests are in analytic (especially multiplicative) and elementary (especially sieve-related) number theory. He is also interested in any part of mathematics which can say something about his favourite few problems, most of which are in some way related to quadratic polynomials.

Andrew Granville, educated at Cambridge University in England and Queen's University in Canada, is now the Canadian research chair in number theory at the Université de Montréal. His main research interests lie in analytic and combinatorial number theory, and in solutions to Diophantine equations. He first became interested in distribution questions when studying primes and smooth numbers. His interest was piqued by Maier's result that primes are not distributed in short intervals as well as the Gauss–Cramér model had suggested.

Roger Heath-Brown was an undergraduate in Cambridge University, where he later did his doctorate under Alan Baker. He moved to Oxford in 1979, and became Professor of Pure Mathematics in 1999. His research has covered a wide range of topics in analytic number theory, including the zeta-function, the distribution of primes, applications of sieves, the circle method, and exponential sums. Most recently he has been particularly interested in the application of analytic methods to the study of Diophantine Geometry.

Elon Lindenstrauss studied mathematics both as an undergraduate and a graduate student at the Hebrew University in Jerusalem, obtaining his Ph.D. under the guidance of Benjamin Weiss. His first postgraduate position was at the Institute for Advanced Study in Princeton where he learnt about arithmetic quantum unique ergodicity and other arithmetic questions from Peter Sarnak. He is currently Professor of Mathematics at Princeton University, and was a Clay Research Fellow for two years. His research interests include ergodic theory, dynamical systems, automorphic forms and number theory and particularly the interplay between these fields.

Jens Marklof is Professor of Mathematical Physics at the University of Bristol. He was educated at the Universities of Hamburg, Princeton and Ulm, where he received his Ph.D. in 1997. Before taking up a lectureship at Bristol in 1999, Marklof held post-doctoral positions at Hewlett-Packard, Cambridge University, Université Paris-Sud and IHÉS. His main interests are problems at the interface of dynamical systems, number theory and quantum mechanics. In 2004 he received the Marie Curie Award and the Philip Leverhulme Prize for his work on the spectral statistics of integrable quantum systems.

Zeév Rudnick received his Ph.D. at Yale University in 1990. He was a Szego assistant professor at Stanford and following that an assistant professor at Princeton University. Since 1995 he has been at Tel Aviv university. Currently, his main research interests are in analytic number theory and in quantum chaos.

K. Soundararajan received his undergraduate degree from the University of Michigan and was a doctoral student of Peter Sarnak at Princeton University. He held a fellowship from the American Institute of Mathematics, and was a postdoc at the Institute for Advanced Study, before moving back to the University of Michigan. From the Fall of 2006 he will be Professor of Mathematics at Stanford University. His main interests are in L-functions and multiplicative number theory, harmonic analysis, and combinatorial number theory.

Yuri Tschinkel was educated at Moscow State University and M.I.T., and is currently (the Gauss) Chair of Pure Mathematics at the University of Göttingen, and is Professor of Mathematics at the Courant Institute, NYU. His research interests lie in arithmetic algebraic geometry and analytic number theory. In particular, he studies the distribution of rational and integral points on higher-dimensional algebraic varieties, over number fields and function fields.

Emmanuel Ullmo is a professor at Université Paris-Sud (Orsay). His main research interests lie in arithmetic geometry (Arakelov theory, modular and automorphic forms, Shimura varieties, ergodic theory). He has worked on several problems of equidistribution of sequences of points or positive dimensional subvarieties of an algebraic variety: points with small heights on abelian varieties, CM points and Hecke points on Shimura varieties, equidistribution of special subvarieties of Shimura varieties.

Akshay Venkatesh was an undergraduate at the University of Western Australia and received his Ph.D. from Princeton University. He is presently an Associate Professor at NYU. His research has focussed on problems with an analytic flavor in number theory and automorphic forms. He became very interested in ergodic theory because "they" kept proving results that he couldn't!

UNIFORM DISTRIBUTION

Andrew Granville
Université de Montréal

Zeév Rudnick
Tel-Aviv University

1. Uniform Distribution mod One

At primary school the first author was taught to estimate the area of a (convex) body by drawing it on a piece of graph paper, and then counting the number of (unit) squares inside. There is obviously a little ambiguity in deciding how to count the squares which straddle the boundary. Whatever the protocol, if the boundary is more-or-less smooth then the number of squares in question is proportional to the perimeter of the body, which will be small compared to the area (if the body is big enough). At secondary school the first author learnt that there are other methods to determine areas, sometimes more precise. As an undergraduate he learned that counting lattice points is often a difficult question (and that counting unit squares is "equivalent" to counting the lattice points in their bottom left-hand corner). Then, as a graduate student, he learnt that the primary school method could be turned around to provide a good tool for estimating the number of lattice points inside a convex body! In the specific case of a right-angled triangle we fix the slope $-\alpha$ of the hypotenuse and ask for the number of lattice points

$$A_\alpha(N) := \#\{(x, y) \in \mathbb{Z}^2 : x, y \geq 0 \text{ and } y + \alpha x \leq N\}.$$

For fixed α the primary school method yields

$$A_\alpha(N) = \frac{N^2}{2\alpha} + O_\alpha(N). \tag{1}$$

Can we improve on the error term $O_\alpha(N)$? For integer m we have

$$\begin{aligned}&\left(A_{-1}(m) - \frac{m^2}{2}\right) - \left(A_{-1}(m - \frac{1}{m}) - \frac{(m - \frac{1}{m})^2}{2}\right)\\ &= A_{-1}(m) - A_{-1}(m-1) + 1 - \frac{1}{2m^2} = \binom{m+1}{2} - \binom{m}{2} + O(1) = m + O(1);\end{aligned}$$

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 1–13.
© 2007 *Springer.*

thus we cannot replace the "$O_{-1}(N)$" term in (1) by "$o_{-1}(N)$." Moreover a similar argument works whenever $\alpha \in \mathbb{Q}$. It is unclear whether (1) can be improved when $\alpha \notin \mathbb{Q}$ so we now examine this case in more detail:

For each integer $x \geq 0$ the number of integers $y \geq 0$ for which $y + \alpha x \leq N$ is simply $\max\{0, 1 + [N - \alpha x]\}$, where $[t]$ denotes the largest integer $\leq t$. Whenever $x \leq N/\alpha$ we can write $1 + [N - \alpha x] = 1 + N - \alpha x - \{N - \alpha x\}$, where $\{t\} = t - [t]$. Therefore

$$\begin{aligned} A_\alpha(N) &= \sum_{x=0}^{[N/\alpha]} (1 + N - \alpha x - \{N - \alpha x\}) \\ &= \frac{N^2}{2\alpha} + \frac{1}{2}\left(N + \frac{N}{\alpha}\right) + O(1) - \sum_{x=0}^{[N/\alpha]} (\{N - \alpha x\} - \frac{1}{2}). \qquad (2) \end{aligned}$$

The first term is indeed the area of our triangle. The second two terms account for half the length of the perimeter of our triangle. So, to able to prove that

$$A_\alpha(N) = \text{Area} + \frac{\text{Perimeter}}{2} + o_\alpha(N),$$

we need to establish that the mean value of $\{N - \alpha x\}$ is $\frac{1}{2}$ when α is irrational, as one might guess. Actually we will prove something much stronger. We will prove that these values, in fact any set of values $\{\alpha n + \beta : 1 \leq n \leq N\}$ with α irrational, are "uniformly distributed mod one," so that their average is $\frac{1}{2}$:

DEFINITION. A sequence of real numbers $a_1, a_2, \ldots$ is *uniformly distributed mod one* if, for all $0 \leq b < c \leq 1$ we have

$$\#\{n \leq N : \ b < \{a_n\} \leq c\} \sim (c - b)N \quad \text{as } N \to \infty.$$

Note that the values $a_n = np/q + \beta$, $1 \leq n \leq N$ (here $\alpha = p/q \in \mathbb{Q}$) are evidently *not* uniformly distributed mod one.

Dirichlet proved that for any integer $M \geq 1$ there exists integer m, $1 \leq m \leq M$ such that $\|m\alpha\| < 1/M$ (where $\|t\|$ denotes the distance from t to the nearest integer). To prove this note that there are $M + 1$ numbers $\{0 \cdot \alpha\}, \{1 \cdot \alpha\}, \ldots, \{M \cdot \alpha\}$ so, by the pigeonhole principle two, say $\{i \cdot \alpha\}$ and $\{j \cdot \alpha\}$ with $0 \leq i < j \leq M$, must belong to the same interval $[k/M, (k + 1)/M)$ and so the result follows with $m = j - i$.

For $\alpha \notin \mathbb{Q}$ we have $\delta := \|m\alpha\| > 0$. We will show that for each i, $1 \leq i \leq m$ the set of values $\{\alpha n + \beta : 1 \leq n \leq N, \ n \equiv i \pmod{m}\}$ is well-distributed mod one, and so the union of these sets is. This set of values is $\{j(m\alpha) + (i\alpha + \beta)\} : 1 \leq j \leq J_i\}$ where $J_i = N/m + O(1)$. We can rewrite this as $\{\delta j+\gamma \pmod 1 : 1 \le j \le J\}$ where $\gamma \equiv i\alpha+\beta \pmod 1$ if $\delta = \{m\alpha\}$, and $\gamma \equiv i\alpha+\beta-\delta(J+1) \pmod 1$ if $1-\delta = \{m\alpha\}$, by replacing j with $J+1-j$. Now, for $0 \le \gamma < 1$ and $K = [\delta J + \gamma]$

$$\begin{aligned}\#\{j \le J : \ \{\delta j+\gamma\} \in [b,c)\} &= \sum_{k=0}^{K} \#\{j \le J : \ \delta j+\gamma \in [k+b, k+c)\} \\ &= (K+O(1))\left(\frac{c-b}{\delta}+O(1)\right) \\ &= (c-b)J + O\left(\frac{c-b}{\delta}+\delta J+1\right).\end{aligned}$$

So fix $\epsilon > 0$ and let $M > 1/\epsilon$ so that $\delta < 1/M < \epsilon$. We have just shown that

$$\#\{n \le N : \{\alpha n+\beta \in [b,c)\}\} = (c-b)N + O\left(\frac{m}{\delta}+\delta N\right).$$

Selecting $N > m/\delta^2$ this is $(c-b+O(\epsilon))N$. Letting $\epsilon \to 0$ we deduce that the sequence $\{\alpha n+\beta : n \ge 1\}$ is uniformly distributed mod one.

The above argument works for linear polynomials in α but it is hard to see how it can be modified for more general sequences. However to determine whether a sequence of real numbers is uniformly distributed we have the following extraordinary, and widely applicable, criterion:

WEYL'S CRITERION ((Weyl, 1914)). *A sequence of real numbers $a_1, a_2, \ldots$ is uniformly distributed mod one* if and only if *for every integer $b \neq 0$ we have*

$$\left|\sum_{n\le N} e(ba_n)\right| = o_b(N) \quad \text{as } N \to \infty. \tag{3}$$

In other words $\limsup_{N\to\infty} \frac{1}{N}|\sum_{n\le N} e(ba_n)| = 0$.

(Here, and throughout, $e(t) := e^{2i\pi t}$.) In particular if $a_n = \alpha n+\beta$ then

$$\sum_{n\le N} e(ba_n) = e(b\beta)\sum_{n\le N} e(b\alpha n) = e(b(\alpha+\beta)) \cdot \frac{e(b\alpha N)-1}{e(b\alpha)-1},$$

the sum of a geometric progression, provided $b\alpha$ is not an integer, so that

$$\left|\sum_{n\le N} e(ba_n)\right| \le \frac{2}{|e(b\alpha)-1|} \asymp \frac{1}{\|b\alpha\|} \ll_b 1 = o_b(N), \tag{4}$$

as $|e(t)-1| \asymp \|t\|$. Since $b\alpha$ is never an integer when $\alpha \notin \mathbb{Q}$ we deduce, from Weyl's criterion, that the sequence $\{\alpha n+\beta : \ n \ge 1\}$ with α irrational, is uniformly distributed mod one.

REMARK. We immediately deduce from Weyl's criterion that if $a_1, a_2, \dots$ is uniformly distributed mod one then so is $ka_1, ka_2, \dots$ for any non-zero integer k. Actually this can be deduced from the definition of uniform distribution mod one.

Proof. We recall that $|\sin t| \leq \|t\|$ so that $|e(t) - 1| \leq \pi\|t\|$.

We begin by assuming that a_1, $a_2, \dots$ is uniformly distributed mod one. Fix integer b and then fix integer $M > b$. Since the sequence is uniformly distributed mod one we know that for each m, $0 \leq m \leq M-1$, there are $N/M + o(N)$ values of $n \leq N$ with $m/M \leq a_n < (m+1)/M$; moreover, for such n, we have $|e(ba_n) - e(bm/M)| \leq \pi\|b/M\|$. Therefore

$$\sum_{n\leq N} e(ba_n) = \sum_{m=0}^{M-1} \left(\frac{N}{M} + o(N)\right)\left(e\left(\frac{bm}{M}\right) + O_b\left(\frac{1}{M}\right)\right) = O_b\left(\frac{N}{M}\right) + o(MN).$$

Now letting M get increasingly large we deduce that our sum is indeed $o_b(N)$.

On the other hand, assume that (1) holds and define the characteristic function $\chi_{(b,c]}$ by $\chi_{(b,c]}(t) = 1$ if $\{t\} \in (b,c]$, and $= 0$ otherwise. A well-known result from Fourier analysis tells us that one can approximate any "reasonable" function arbitrarily well using polynomials. That is, for any $\epsilon > 0$ there exists integer d and coefficients c_j, $-d \leq j \leq d$, such that we have $\left|\chi(t) - f(e(t))\right| \leq \epsilon$ for all $t \in [0,1)$ where $f(x) = \sum_{j:|j|\leq d} c_j x^j$. Therefore

$$\begin{aligned}\#\{n \leq N : b < \{a_n\} \leq c\} &= \sum_{n\leq N} \chi_{(b,c]}(a_n) = \sum_{n\leq N} \Big(f(e(a_n)) + O(\epsilon)\Big) \\ &= \sum_{j:|j|\leq d} c_j \sum_{n\leq N} e(ja_n) + O(N\epsilon) = c_0 N + o(N) + O(N\epsilon)\end{aligned}$$

by (4). Now

$$c - b = \int_0^1 \chi_{(b,c]}(t)dt = \sum_{j:|j|\leq d} c_j \int_0^1 e(jt)\,dt + O(\epsilon) = c_0 + O(\epsilon)$$

and so, by combining the last two equations and letting $\epsilon \to 0$, we have shown that the sequence is uniformly distributed mod one. □

One can deduce that $a_1, a_2, \dots$ is uniformly distributed mod one if and only if, for every continuous function f: $[0,1) \to \mathbb{R}$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n\leq N} f(\{a_n\}) = \int_0^1 f(x)\,dx.$$

To prove this note that the functions $e(bx)$, $b \in \mathbb{Z}$ form an appropriate (Fourier) basis for the continuous functions on $[0,1)$.

An explicit version of Weyl's result, which is useful for many applications, was given by Erdős and Turán (Erdős and Turán, 1948): For any sequence of real numbers, and any $0 \le b < c \le 1$ we have

$$\left|\frac{1}{N}\#\{n \le N : \ b < \{a_n\} \le c\} - (c-b)\right| \le \frac{6}{m+1} + \frac{4}{\pi}\sum_{b=1}^{m}\frac{1}{b}\left|\frac{1}{N}\sum_{n\le N} e(ba_n)\right|.$$

There is a nice application of Weyl's theorem in the theory of elliptic curves: Let E be an elliptic curve defined over $\mathbb{Q}$ and suppose that E has infinitely many rational points. Poincaré showed that the rational points form an additive group, and Mordell proved Poincaré's conjecture that this group has finite rank; in other words $E(\mathbb{Q})$ is an additive group of the form $\mathbb{Z}^r \bigoplus T$ where the torsion subgroup T (that is, the subgroup of points of finite order) and r are finite. Let us suppose that $P_1, \ldots, P_r$ form a basis for the $\mathbb{Z}^r$ part of $E(\mathbb{Q})$: For any given arc A on $E(\mathbb{R})$ we can ask what proportion of the points $\{n_1P_1 + n_2P_2 + \ldots + n_rP_r + t : 0 \le n_1, \ldots, n_r \le N-1, \ t \in T\}$ lie on A, as $N \to \infty$? The connection with our work above lies in the Weierstrass parameterization of E: There exists an isomorphism $\wp\colon \mathbb{C}/(\mathbb{Z}+\mathbb{Z}i) \to E$; that is $\wp(v+w) = \wp(v) + \wp(w)$ for all $v, w \in \mathbb{C}$. So select $z_1, \ldots, z_r \in \mathbb{C}$ such that $\wp(z_j) = P_j$ and τ such that $\wp(\tau) = t$. The above question then becomes to determine the proportion of the points

$$\{n_1z_1+n_2z_2+\cdots+n_rz_r+\tau \pmod{\mathbb{Z}+\mathbb{Z}i} : 0 \le n_1, \ldots, n_r \le N-1, \ \tau \in \wp^{-1}(T)\}$$

that lie on $\wp^{-1}(A)$, a two-dimensional uniform distribution question. Like this the proportion can be shown to be

$$\int_{(x,y)\in A} \frac{dx}{y} \Big/ \int_{(x,y)\in E(\mathbb{R})} \frac{dx}{y}.$$

(For more background on elliptic curves see (Silverman and Tate, 1992)).

For given $v = (a_1, \ldots, a_k) \in \mathbb{R}^k$ define $v \pmod 1$ to be the vector $(a_1 \pmod 1), \ldots, a_k \pmod 1))$. We say that the sequence of vectors $v_1, v_2, \ldots \in \mathbb{R}^k$ is *uniformly distributed mod one* if for any $0 \le b_j < c_j < 1$ for $j = 1, 2, \ldots, k$, we have

$$\#\left\{n \le N : a_n \pmod 1 \in \bigoplus_{j=1}^{k}[b_j, c_j)\right\} \sim \prod_{j=1}^{k}(c_j - b_j)\cdot N \quad \text{as } N \to \infty.$$

WEYL'S CRITERION IN K DIMENSIONS. *A sequence of vectors $v_1, v_2, \ldots \in \mathbb{R}^k$ is uniformly distributed mod one* if and only if *for every $b \in \mathbb{Z}^k$, $b \ne 0$ we have*

$$\left|\sum_{n\le N} e(b.v_n)\right| = o_b(N) \quad \textit{as } N \to \infty. \tag{5}$$

We can deduce Kronecker's famous result that if $1, \alpha_1, \alpha_2, \dots, \alpha_k$ are linearly independent over $\mathbb{Q}$ then the vectors $\{(n\alpha_1, n\alpha_2, \dots, n\alpha_k) : n \geq 1\}$ are uniformly distributed mod one.

A final remark on $\{\alpha n + \beta\}_{n\geq 1}$: Let $a_n = \alpha n + \beta \pmod 1$ for all $n \geq 1$. The transformation $T_\alpha : x \to x + \alpha$ gives $T : a_n \to a_{n+1}$. We want to define a measure μ on $\mathbb{R}/\mathbb{Z}$ such that, for any "sensible" set A we have $\mu(A) = \mu(T_\alpha^{-1}A)$. In fact, when $\alpha \notin \mathbb{Q}$, the only invariant such measure, μ, is the Lebesgue measure, and thus the values a_n are distributed according to this measure, that is they are uniformly distributed mod one. See Section 2.4 of Lindenstrauss's paper in this volume (Lindenstrauss, 2006) for more details of this kind of ergodic theoretic proof.

2. Fractional Parts of αn^2

We have seen, in the last section, that any sequence $\{\alpha n + \beta : n \geq 1\}$, with α irrational, is uniformly distributed mod one. One might ask about higher degree polynomials in n. Our goal in this section is to prove the following celebrated theorem of H. Weyl:

THEOREM 2.1. *For any irrational real number α, the sequence $\{\alpha n^2 : n \geq 1\}$ is uniformly distributed mod one.*

For a streamlined proof, see the book (Kuipers and Niederreiter, 1974). Here we will give an argument close to the original:

By Weyl's criterion, we need to show that for fixed integer $b \neq 0$, the "Weyl sum"

$$S_\beta(N) = \sum_{n=1}^{N} e(\beta n^2)$$

is $o_\beta(N)$, where $\beta = b\alpha$. Note that β is also irrational.

Weyl's idea was to *square* the sum and notice that the resulting sum is essentially that of a polynomial one degree lower, that is a linear polynomial. Indeed,

$$|S_\beta(N)|^2 = \sum_{x,y\leq N} e(\beta(x^2 - y^2)) = N + 2\Re \sum_{y>x} e(\beta(y^2 - x^2))$$

Writing $y = x + h$, with $h = 1, \dots, N-1$, $x = 1, \dots, N-h$ we have $y^2 - x^2 = 2hx + h^2$ which is *linear* in x. Thus we find

$$|S_\beta(N)|^2 \;=\; N + 2\Re \sum_{h=1}^{N-1} e(\beta h^2) \sum_{x=1}^{N-h} e(2\beta h \cdot x)$$

$$\begin{aligned} &\leq N + 2\sum_{h=1}^{N-1}\left|\sum_{x=1}^{N-h} e(2\beta h\cdot x)\right| \\ &\ll N + \sum_{h=1}^{N-1}\min\left\{N, \frac{1}{\|2\beta h\|}\right\}, \end{aligned} \tag{6}$$

proceeding as in (4).

We again use Dirichlet's observation that there exists $q \leq N$ with $\|q(2\beta)\| < 1/N$. Let a be the integer nearest $q(2\beta)$; we may assume that $(a, q) = 1$. If $h = H + j$, $1 \leq j \leq q$ then $\|2\beta h\| = \|2\beta H + aj/q\| + O(1/N)$; so as j runs from 1 to q the values $\|2\beta h\|$ (where $h = H + j$) run through the values $\|\gamma + i/q\|$ for $0 \leq i \leq q-1$, with error no more than $O(1/N)$, where $|\gamma| \leq 1/2q$. Thus,

$$\sum_{h=H+1}^{H+q}\min\left\{N, \frac{1}{\|2\beta h\|}\right\} \ll N + \sum_{i=1}^{q/2}\frac{q}{i} \ll N + q\log q.$$

Partitioning the integers up to $N-1$ into at most $N/q + 1 \leq 2N/q$ intervals of length q or less, we thus deduce, from (6), that

$$\left|\frac{1}{N}S_\beta(N)\right|^2 \ll \frac{1}{q} + \frac{\log q}{N}. \tag{7}$$

Now $q = q_N \to \infty$ as $N \to \infty$ so (7) is $o(1)$ and we are done. To see that $q_N \to \infty$ as $N \to \infty$, suppose not so that $\|q(2\beta)\| < \frac{1}{N}$ for infinitely many integers N and thus $\|q(2\beta)\| = 0$. But then β can be written as a rational number with denominator $2q$, contradicting hypothesis.

This result is widely applicable and this proof is easily modified to fit a given situation. For example see the proof of Lemma 3.2 in Heath-Brown's paper (Heath-Brown, 2006) in this volume.

A rather elegant ergodic theoretic proof of Theorem 2.1 is given in Section 3 of Lindenstrauss's paper in this volume (Lindenstrauss, 2006).

Theorem 2.1 is a special case of

THEOREM 2.2. *Let $P(x) = a_dx^d + a_{d-1}x^{d-1} + \cdots + a_1x + a_0$ be a polynomial with at least one of the coefficients $a_1, \ldots a_d$ irrational. Then the sequence $\{P(n) : n \geq 1\}$ is uniformly distributed modulo* 1.

This can be proved along the same lines as Theorem 2.1 (the special case of the polynomial $P(x) = \alpha x^2$) except that a single squaring operation will now produce a polynomial of one degree less, not a linear one. One then iterates this procedure to get back to the case of linear polynomials, see, e.g., (Davenport, 2005) for details.

One can deduce from Weyl's criterion in n-dimension and Theorem 2.2 that the vectors $\{(n\alpha, n^2\alpha, \dots, n^k\alpha) \colon n \geq 1\}$ are uniformly distributed mod one if α is not rational (see also Lindenstrauss's article (Lindenstrauss, 2006) in this volume).

3. Uniform Distribution mod N

For a given set A define $A(x) = 1$ if $x \in A$, and $A(x) = 0$ otherwise. Also define the *Fourier transform of A* to be

$$\hat{A}(b) := \sum_n A(n)e(bn) = \sum_{n \in A} e(bn).$$

Writing $A_N = \{a_j : 1 \leq j \leq N\}$ the Weyl criterion becomes that $a_1, a_2, \dots$ is uniformly distributed mod one if and only if $\hat{A}_N(b) = o_b(N)$ for every non-zero integer b.

When A is a subset of the residues mod N we define

$$\hat{A}(b) := \sum_n A(n)e\left(\frac{bn}{N}\right) = \sum_{n \in A} e\left(\frac{bn}{N}\right).$$

Let A be a set of integers, and let $(t)_N$ denote the least non-negative residue of $t \pmod N$ (so that $(t)_N = N\{t/N\}$). The idea of uniform distribution mod N is surely something like: For all $0 \leq b < c \leq 1$ and all $m \not\equiv 0 \pmod N$ we have

$$\#\{a \in A : \ bN < (ma)_N \leq cN\} \sim (c-b)|A|. \tag{8}$$

One can only make such a definition if $|A| \to \infty$ (since this is an asymptotic formula) but we are often interested in smaller sets A, indeed that are a subset of $\{1, 2, \dots, N\}$; so we will work with something motivated by, but different from, (8). Let us see how far we can go to proving the analogy to Weyl's criterion. Fix $\epsilon > 0$:

Define

$$\text{Error}(A;k) := \max_{\substack{0 \leq x \leq N \\ m \not\equiv 0 \pmod N}} \left| \# \left\{ a \in A : x < (ma)_N \leq x + \frac{N}{k} \right\} - \frac{|A|}{k} \right|.$$

Suppose that $\text{Error}(A;k) \leq \epsilon|A|/k$ for some $k > 1/\epsilon$ We proceed much as in the proof of Weyl's criterion above: Subdivide our interval $(0, N]$ into subintervals $I_j := (jN/k, (j+1)N/k]$, so that if $(ma)_N \in I_j$ then $e(ma/N) = e(j/k) + O(1/k)$. Therefore

$$\hat{A}(m) = \sum_{j=0}^{k-1} \sum_{\substack{a \in A \\ (ma)_N \in I_j}} e(ma/N) = \sum_{j=0}^{k-1} e(j/k) \sum_{\substack{a \in A \\ (ma)_N \in I_j}} 1 + O(|A|/k)$$

$$\ll\ k\,\mathrm{Error}(A;k)+\frac{|A|}{k}\ll\epsilon|A|.$$

In the other direction our proof is somewhat different from that for Weyl's criterion: We begin by supposing that $|\hat{A}(b)|\leq\epsilon^2|A|$ for all $b\not\equiv 0 \pmod N$. For $J=[\delta N]$

$$\sum_{\substack{a\in A\\ 1\leq(ma)_N\leq J}} 1=\sum_{j=1}^{J}\sum_{a\in A}\frac{1}{N}\sum_{r}e\left(r\left(\frac{ma-j}{N}\right)\right)=\frac{J}{N}|A|+\frac{1}{N}\sum_{r\neq 0}\hat{A}(rm)\sum_{j=1}^{J}e\left(\frac{-rj}{N}\right).$$

If r runs through the non-zero integers in $(-N/2,N/2]$ then $|\sum_{j=1}^{J}e(-rj/N)|\ll N/|r|$. Thus the second term here is, for $R\approx N/(\epsilon^2|A|)$

$$\begin{aligned}
&\ll \sum_{r\neq 0}\frac{|\hat{A}(rm)|}{r}\leq\sum_{0\leq|r|\leq R}\frac{|\hat{A}(rm)|}{r}+\sum_{R<|r|\leq N/2}\frac{|\hat{A}(rm)|}{r}\\
&\leq (\log R)\max_{s\neq 0}|\hat{A}(s)|+\left(\sum_{r}|\hat{A}(rm)|^2\right)^{1/2}\left(\sum_{R<|r|}1/r^2\right)^{1/2}\\
&\leq (\log R)\epsilon^2|A|+(|A|N/R)^{1/2}\ll\epsilon|A|
\end{aligned}$$

provided $\epsilon\ll 1/\log(N/|A|)$.

One can thus formulate an appropriate analogy to Weyl's criterion along the lines: The Fourier transforms of A are all small if and only if A and all its dilates are "uniformly distributed." (A *dilate* of A is the set $\{ma : a\in A\}$ for some $m\not\equiv 0 \pmod N$.) This result is central to the spectacular recent progress in harmonic analysis by Gowers et al., (see (Granville et al., 2006)).

To give one example of how such a notion can be used, we ask whether a given set A of residues mod N contains a non-trivial 3-term arithmetic progression? In other words we wish to find solutions to $a+b=2c$ with $a,b,c\in A$ where $a\neq b$.

PROPOSITION 3.1. *If A is a subset of the residues $\pmod N$ where N is odd, for which $|\hat{A}(m)|<|A|^2/N-1$ whenever $m\not\equiv 0 \pmod N$ then A contains non-trivial 3-term arithmetic progressions.*

Proof. Since $(1/N)\sum_r e(rt/N)=0$ unless t is divisible by N, whence it equals 1, we have that the number of 3-term arithmetic progressions in A is

$$\sum_{a,b,c\in A}\frac{1}{N}\sum_{r}e\left(\frac{r(a+b-2c)}{N}\right)=\frac{1}{N}\sum_{r}\hat{A}(r)^2\hat{A}(-2r).$$

The $r=0$ term gives $|A|^3/N$. We regard the remaining terms as error terms, and bound them by their absolute values, giving a contribution (by taking

$m \equiv -2r \pmod{N})$)

$$\leq \frac{1}{N}\sum_r |\hat{A}(r)|^2 \cdot \max_{m\neq 0}|\hat{A}(m)| = |A|\max_{m\neq 0}|\hat{A}(m)|.$$

There are $|A|$ trivial 3-term arithmetic progressions (of the form a, a, a) so we have established that A has non-trivial 3-term arithmetic progressions when

$$|A|^3/N - |A|\max_{m\neq 0}|\hat{A}(m)| > |A|,$$

yielding the result.

Let us apply Proposition 3.1 to the sets

$$A_\delta := \left\{ n \pmod{N} : \left\| \frac{n^2}{N} \right\| < \frac{\delta}{2} \right\}$$

for N prime with $0 < \delta < 1$. For $J = [\delta N/2]$ we have

$$\hat{A}_\delta(m) = \sum_n e\left(\frac{mn}{N}\right) \sum_{-J\leq j\leq J} \frac{1}{N} \sum_r e\left(r\frac{(j-n^2)}{N}\right)$$

so that

$$|\hat{A}_\delta(m)| \leq \frac{1}{N}\sum_r \left| \sum_{-J\leq j\leq J} e\left(\frac{rj}{N}\right)\right| \left|\sum_n e\left(\frac{mn-rn^2}{N}\right)\right|.$$

Now $\sum_n e(mn/N) = 0$ if $m \neq 0$, and $= N$ if $m = 0$. If $r \neq 0$ then $\sum_n e((mn - rn^2)/N)$ is a Gauss sum and so has absolute value $\sqrt{N}$. Moreover $|\sum_{-J\leq j\leq J} e(rj/N)| \ll N/r$ for $1 \leq r \leq N/2$. Inputting all this above we obtain $|\hat{A}_\delta(m)| \ll \sqrt{N}\log N$ for each $m \not\equiv 0 \pmod{N}$ and $\#A_\delta = |\hat{A}_\delta(0)| = \delta N + O(\sqrt{N}\log N)$. Now, for fixed $\delta > 0$ we have proved that each $|\hat{A}_\delta(m)| = o(\delta^2 N)$, and so Proposition 3.1 implies that A_δ contains non-trivial 3-term arithmetic progressions. In fact the proof of Proposition 3.1 yields that A_δ has $\sim \delta^3 N^2$ 3-term arithmetic progressions $a, a+d, a+2d$.

The previous result is in fact a special case of Roth's (Roth, 1953) theorem, which states that for any $\delta > 0$ if N is sufficiently large then any subset A of $\{1, \ldots, N\}$ with more than δN elements contains a non-trivial 3-term arithmetic progression. His proof is a little too complicated to discuss in detail here but we will outline the main ideas. If $\delta > \frac{2}{3}$ then A must contain three consecutive integers, so the result follows. Otherwise we proceed by a form of induction, showing that if there exists $A \subset \{1, \ldots, N\}$, with $\#A \sim \delta N$, which contains no non-trivial 3-term arithmetic progression then there exists $A' \subset \{1, \ldots, N'\}$, with $\#A' \sim \delta' N'$, which contains no non-trivial 3-term arithmetic progression, where $\delta' = (1 + c\delta)\delta$ and $N' = [N^{1/3}]$. The induction then yields Roth's theorem for any $\delta \gg 1/\log\log N$. To prove the induction step we begin by increasing N by a negligible amount so that it is prime, and then considering A as a set of residues mod N. By a slight modification of the proof of Proposition 3.1 one can show that if A does not contain a non-trivial 3-term arithmetic progression then A is not uniformly distributed mod N. By the definition of uniformly distributed mod N, this implies that there is some dilate of A, say $mA \pmod N$ and some segment $[bN, cN]$ which contains rather more or rather less elements than expected; one can show that, in fact, there must be some segments with rather more, and some segments with rather less. Taking one of these segments with rather more elements than expected, in fact containing $1 + c\delta$ times as many elements as expected, we can identify a segment of an arithmetic progression (of length N') within $\{1, \dots, N\}$ which contains $\sim \delta' N'$ elements of A, and from this we construct A' (integer $j \in A'$ if and only if the jth term of the arithmetic progression is in A).

4. Normal Numbers

Are there any patterns in the digits of π? Science fiction writers (Sagan, 1985) would have us believe that secret messages are encoded far off in the tail of π but computational evidence so far suggests the contrary, that there are no patterns, indeed that every sequence of digits appears about as often as in a random sequence. If the digits are written in base 10 then this question is equivalent to asking whether the sequence $\{10^n\pi : n \geq 1\}$ is uniformly distributed mod one? If so we say that π *is normal in base* 10. In general we say that real number α *is normal in base* b if the sequence $\{b^n\alpha : n \geq 1\}$ is uniformly distributed mod one; and that α *is normal*, if it is normal in base b for every integer $b \geq 2$.

In general very little is known about normality. A few specific numbers of very special form can be shown to be normal to certain bases. The one thing that we can show is that almost all numbers are normal, with a proof that fails to identify any such number!

THEOREM 4.1. *Almost all* $x \in [0, 1)$ *are normal.*

(By "almost all" we mean that the set of such x has measure 1.) Theorem 4.1 follows from:

THEOREM 4.2. *For any increasing sequence of integers* $a_1, a_2, \dots$*, the sequence* $\{a_n x : n \geq 1\}$ *is uniformly distributed mod one for almost all* $x \in [0, 1)$.

Deduction of Theorem 4.1. Taking $a_j = b^j$ for each j we see that almost all $x \in [0, 1)$ are normal in base b. Theorem 4.1 follows since the exceptional set has measure 0 as it is a countable union of measure 0 sets.

Proof of Theorem 4.2. We begin by noting that

$$\int_0^1 \left| \frac{1}{N} \sum_{n \le N} e(ba_n x) \right|^2 dx = \frac{1}{N^2} \sum_{m,n \le N} \int_0^1 e(bx(a_m - a_n))\, dx = \frac{1}{N};$$

so that

$$\int_0^1 \sum_{m \ge 1} \left| \frac{1}{m^2} \sum_{n \le m^2} e(ba_n x) \right|^2 dx = \sum_{m \ge 1} \frac{1}{m^2} = \frac{\pi^2}{6}.$$

Therefore (in a step that takes some thinking about)

$$\sum_{m \ge 1} \left| \frac{1}{m^2} \sum_{n \le m^2} e(ba_n x) \right|^2 < \infty$$

for almost all x, and so

$$\lim_{m \to \infty} \left| \frac{1}{m^2} \sum_{n \le m^2} e(ba_n x) \right| = 0.$$

Now if $m^2 \le N < (m+1)^2$ then $\sum_{n \le N} e(ba_n x) = \sum_{n \le m^2} e(ba_n x) + O(m)$ and the result follows.

References

Davenport, H. (2005) *Analytic methods for Diophantine equations and Diophantine inequalities*, Cambridge, Cambridge University Press.

Erdős, P. and Turán, P. (1948) On a problem in the theory of uniform distribution I, II, *Indag. Math.* **10**, 370–378, 406–413.

Granville, A., Nathanson, M., and Solymosi, J. (eds.) (2006) *Additive Combinatorics, a school and workshop*, Providence, RI, Amer. Math. Soc., to appear.

Heath-Brown, D. R. (2006) Analytic methods for the distribution of rational points on algebraic varieties, in this volume.

Kuipers, L. and Niederreiter, H. (1974) *Uniform distribution of sequences*, Pure and Applied Mathematics, New York–London–Sydney, Wiley-Interscience.

Lindenstrauss, E. (2006) Three examples of how to use measure classification in number theory, in this volume.

Roth, K. F. (1953) On certain sets of integers, *J. London Math. Soc.* **28**, 104–109.

Sagan, C. (1985) *Contact: A Novel*, New York, Simon and Schuster.

Silverman, J. and Tate, J. (1992) *Introduction to elliptic curves*, New York, Springer-Verlag.

Weyl, H. (1914) Über ein Problem aus dem Gebeit der diophantischen Approximationen, *Nachr. Ges. Wiss. Göttingen (math.-phys. Kl.)* pp. 234–244.

SIEVING AND THE ERDŐS–KAC THEOREM

Andrew Granville
Université de Montréal

K. Soundararajan
University of Michigan

Abstract. We give a relatively easy proof of the Erdős-Kac theorem via computing moments. We show how this proof extends naturally in a sieve theory context, and how it leads to several related results in the literature.

Let $\omega(n)$ denote the number of distinct prime factors of the natural number n. The average value of $\omega(n)$ as n ranges over the integers below x is

$$\frac{1}{x}\sum_{n\le x}\omega(n)=\frac{1}{x}\sum_{p\le x}\sum_{\substack{n\le x\\ p|n}}1=\frac{1}{x}\sum_{p\le x}\left[\frac{x}{p}\right]=\frac{1}{x}\sum_{p\le x}\left(\frac{x}{p}+O(1)\right)=\log\log x+O(1).$$

It is natural to ask how $\omega(n)$ is distributed as one varies over the integers $n\le x$. A famous result of Hardy and Ramanujan (Hardy and Ramanujan, 1917) tells us that $\omega(n)\sim\log\log x$ for almost all $n\le x$; we say that $\omega(n)$ has *normal order* $\log\log n$. To avoid confusion let us state this precisely: given $\epsilon>0$ there exists x_ϵ such that if $x\ge x_\epsilon$ is sufficiently large, then $(1+\epsilon)\log\log x\ge\omega(n)\ge(1-\epsilon)\log\log x$ for all but at most ϵx integers $n\le x$. The functions $\log\log n$ and $\log\log x$ are interchangeable here since they are very close in value for all but the tiny integers $n\le x$.

Their proof revolves around the following wonderful inequality which they established by induction. Define $\pi_k(x)$ to be the number of integers $n\le x$ with $\omega(n)=k$. There exist constants $c_0,c_1>0$ such that for any $k\ge 0$ we have

$$\pi_k(x)<c_0\frac{x}{\log x}\frac{(\log\log x+c_1)^{k-1}}{(k-1)!},\tag{1}$$

for all $x\ge 2$. Hardy and Ramanujan exploited this by deducing that

$$\sum_{|k-\log\log x|\ge\epsilon\log\log x}\pi_k(x)\le c_0\frac{x}{\log x}\sum_{|k-\log\log x|\ge\epsilon\log\log x}\frac{(\log\log x\ +c_1)^{k-1}}{(k-1)!},$$

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 15–27.

which is easily shown to be about $x/(\log x)^\alpha$ where $\alpha = \alpha_\epsilon = \epsilon^2/2 + O(\epsilon^3)$, far less than ϵx. In fact Hardy and Ramanujan squeezed a little more out of this idea, showing that if $\kappa(n) \to \infty$ as $n \to \infty$, no matter how slowly, then

$$|\omega(n) - \log\log n| \le \kappa(n)\sqrt{\log\log n} \tag{2}$$

for almost all integers $n \le x$.

Once we know that $\omega(n)$ has normal order $\log\log n$, we can ask finer questions about the distribution of $\omega(n)$. For instance how is $\omega(n) - \log\log n$ distributed? More specifically, how big is this typically in absolute value? Turán (Turán, 1934) found a very simple proof of the Hardy–Ramanujan result by showing that

$$\frac{1}{x}\sum_{n\le x}(\omega(n) - \log\log n)^2 = \{1 + o(1)\}\log\log x. \tag{3}$$

One deduces easily that $\omega(n)$ has *normal order* $\log\log n$: For, if there are $m_\epsilon(x)$ integers $\le x$ for which $|\omega(n) - \log\log n| \ge \epsilon\log\log x$ then by (3), $m_\epsilon(x) \le (1/\epsilon^2 + o(1))x/\log\log x$, which is $\le \epsilon x$ for sufficiently large x. Indeed the same argument also gives (2) for almost all $n \le x$.

We have now obtained some information about the distribution of $\omega(n)$, its average value, and the average difference between the value and the mean. Next we ask whether there is a distribution function for $\omega(n)$? In other words if, typically, the distance between $\omega(n)$ and $\log\log n$ is roughly of size $\sqrt{\log\log n}$ can we say anything about the distribution of

$$\frac{\omega(n) - \log\log n}{\sqrt{\log\log n}}? \tag{4}$$

In the late 1930s Mark Kac noticed that these developments bore more than a passing resemblance to developments in probability theory. He suggested that perhaps this distribution is *normal* and even conjectured certain number theory estimates which would imply that. Soon after describing this in a lecture, at which Paul Erdős was in the audience, Erdős and Kac were able to announce the result (Erdős and Kac, 1940): For any $\tau \in \mathbb{R}$, the proportion of the integers $n \le x$ for which $\omega(n) \le \log\log n + \tau\sqrt{\log\log n}$ tends to the limit

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\tau} \mathrm{e}^{-t^2/2}dt \tag{5}$$

as $x \to \infty$. In other words the quantity in (4) is distributed like a normal distribution with mean 0 and variance 1.

Erdős and Kac's original proof was based on the central limit theorem, and Brun's sieve. A different proof follows from the work of Selberg (Selberg,

1954) (extending and simplifying the work of (Sathe, 1953)) who obtained an asymptotic formula for $\pi_k(x)$ uniformly in a wide range of k. Yet a third proof is provided by Halberstam (Halberstam, 1955) who showed how to compute the moments

$$\sum_{n\leq x}(\omega(n) - \log\log x)^k, \tag{6}$$

for natural numbers k, and showed that these agreed with the moments of a normal distribution. Since the normal distribution is well-known to be determined by its moments, he deduced the Erdős-Kac theorem.

In this article, we give a simple method to compute the moments (6), and in fact we can obtain an asymptotic formula uniformly in a wide range of k. Then we discuss how such moments can be formulated for more general sequences assuming sieve type hypotheses.

THEOREM 1. *For any natural number k we let $C_k = \Gamma(k+1)/(2^{k/2}\Gamma(k/2+1))$. Uniformly for even natural numbers $k \leq (\log\log x)^{1/3}$ we have*

$$\sum_{n\leq x}(\omega(n) - \log\log x)^k = C_k x(\log\log x)^{k/2}\left(1 + O\left(\frac{k^{3/2}}{\sqrt{\log\log x}}\right)\right),$$

and uniformly for odd natural numbers $k \leq (\log\log x)^{1/3}$ we have

$$\sum_{n\leq x}(\omega(n) - \log\log x)^k \ll C_k x(\log\log x)^{k/2}\frac{k^{3/2}}{\sqrt{\log\log x}}.$$

We will deduce this theorem from the following technical proposition.

PROPOSITION 2. *Define*

$$f_p(n) = \begin{cases} 1 - \frac{1}{p} & \text{if } p \mid n \\ -\frac{1}{p} & \text{if } p \nmid n. \end{cases}$$

Let $z \geq 10^6$ be a real number. Uniformly for even natural numbers $k \leq (\log\log z)^{\frac{1}{3}}$ we have

$$\sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^k = C_k x(\log\log z)^{k/2}\left(1 + O\left(\frac{k^3}{\log\log z}\right)\right) + O(2^k\pi(z)^k), \tag{7}$$

while, uniformly for odd natural numbers $k \leq (\log\log z)^{1/3}$, we have

$$\sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^k \ll C_k x(\log\log z)^{k/2}\frac{k^{3/2}}{\sqrt{\log\log z}} + 2^k\pi(z)^k. \tag{8}$$

Deduction of Theorem 1. We seek to evaluate $\sum_{n\le x}(\omega(n)-\log\log x)^k$ for natural numbers $k\le(\log\log x)^{1/3}$. Set $z=x^{1/k}$ and note that, for $n\le x$,

$$\omega(n)-\log\log x=\sum_{p\le z}f_p(n)+\sum_{\substack{p|n\\ p>z}}1+\Big(\sum_{p\le z}1/p-\log\log x\Big)=\sum_{p\le z}f_p(n)+O(k).$$

Thus for some positive constant c we obtain that

$$(\omega(n)-\log\log x)^k=\Big(\sum_{p\le z}f_p(n)\Big)^k+O\Big(\sum_{\ell=0}^{k-1}(ck)^{k-\ell}\binom{k}{\ell}\Big|\sum_{p\le z}f_p(n)\Big|^\ell\Big).$$

When we sum this up over all integers $n\le x$ the first term above is handled through (7, 8). To handle the remainder terms we estimate $\sum_{n\le x}\big|\sum_{p\le z}f_p(n)\big|^\ell$ for $\ell\le k-1$. When ℓ is even this is once again available through (7). Suppose ℓ is odd. By Cauchy–Schwarz we get that

$$\sum_{n\le x}\Big|\sum_{p\le z}f_p(n)\Big|^\ell\le\Big(\sum_{n\le x}\Big(\sum_{p\le z}f_p(n)\Big)^{\ell-1}\Big)^{1/2}\Big(\sum_{n\le x}\Big(\sum_{p\le z}f_p(n)\Big)^{\ell+1}\Big)^{1/2},$$

and using (7) we deduce that this is

$$\ll\sqrt{C_{\ell-1}C_{\ell+1}}x(\log\log z)^{\ell/2}.$$

Proof of Proposition 2. If $r=\prod_i p_i^{\alpha_i}$ is the prime factorization of r we put $f_r(n)=\prod_i f_{p_i}(n)^{\alpha_i}$. Then we may write

$$\sum_{n\le x}\Big(\sum_{p\le z}f_p(n)\Big)^k=\sum_{p_1,\dots,p_k\le z}\sum_{n\le x}f_{p_1\cdots p_k}(n).$$

To proceed further, let us consider more generally $\sum_{n\le x}f_r(n)$.

Suppose $r=\prod_{i=1}^s q_i^{\alpha_i}$ where the q_i are distinct primes and $\alpha_i\ge1$. Set $R=\prod_{i=1}^s q_i$ and observe that if $d=(n,R)$ then $f_r(n)=f_r(d)$. Therefore, with τ denoting the divisor function,

$$\begin{aligned}\sum_{n\le x}f_r(n)&=\sum_{d|R}f_r(d)\sum_{\substack{n\le x\\ (n,R)=d}}1=\sum_{d|R}f_r(d)\Big(\frac{x}{d}\frac{\varphi(R/d)}{R/d}+O(\tau(R/d))\Big)\\&=\frac{x}{R}\sum_{d|R}f_r(d)\varphi(R/d)+O(\tau(R)).\end{aligned}$$

Thus setting

$$G(r):=\frac{1}{R}\sum_{d|R}f_r(d)\varphi(R/d)=\prod_{q^\alpha\|r}\Big(\frac{1}{q}\Big(1-\frac{1}{q}\Big)^\alpha+\Big(\frac{-1}{q}\Big)^\alpha\Big(1-\frac{1}{q}\Big)\Big),$$

we conclude that

$$\sum_{n\le x} f_r(n) = G(r)x + O(\tau(R)).$$

Observe that $G(r) = 0$ unless r is square-full and so

$$\sum_{n\le x}\Big(\sum_{p\le z} f_p(n)\Big)^k = x \sum_{\substack{p_1,\dots,p_k\le z\\ p_1\cdots p_k \text{ square-full}}} G(p_1\cdots p_k) + O(2^k\pi(z)^k). \tag{9}$$

Suppose $q_1 < q_2 < \ldots < q_s$ are the distinct primes in $p_1\cdots p_k$. Note that since $p_1\cdots p_k$ is square-full we have $s \le k/2$. Thus our main term above is

$$\sum_{s\le k/2}\ \sum_{q_1<q_2<\ldots<q_s\le z}\ \sum_{\substack{\alpha_1,\dots,\alpha_s\ge 2\\ \sum_i \alpha_i = k}} \frac{k!}{\alpha_1!\cdots\alpha_s!} G(q_1^{\alpha_1}\cdots q_s^{\alpha_s}).$$

When k is even there is a term $s = k/2$ (and all $\alpha_i = 2$) which gives rise to the Gaussian moments. This term contributes

$$\frac{k!}{2^{k/2}(k/2)!} \sum_{\substack{q_1,\dots,q_{k/2}\le z\\ q_i \text{ distinct}}} \prod_{i=1}^{k/2} \frac{1}{q_i}\Big(1-\frac{1}{q_i}\Big).$$

By ignoring the distinctness condition, we see that the sum over q's is bounded above by $(\sum_{p\le z}(1-1/p)/p)^{k/2}$. On the other hand, if we consider $q_1,\dots,q_{k/2-1}$ as given then the sum over $q_{k/2}$ is plainly at least $\sum_{\pi_{k/2}\le p\le z}(1-1/p)/p$ where we let π_n denote the nth smallest prime. Repeating this argument, the sum over the q's is bounded below by $(\sum_{\pi_{k/2}\le p\le z}(1-1/p)/p)^{k/2}$. Therefore the term with $s = k/2$ contributes

$$\frac{k!}{(k/2)!2^{k/2}}\big(\log\log z + O(1+\log\log k)\big)^{k/2}. \tag{10}$$

To estimate the terms $s < k/2$ we use that $0 \le G(q_1^{\alpha_1}\cdots q_s^{\alpha_s}) \le 1/(q_1\cdots q_s)$ and so these terms contribute

$$\le \sum_{s<k/2} \frac{k!}{s!}\Big(\sum_{q\le z}\frac{1}{q}\Big)^s \sum_{\substack{\alpha_1,\dots,\alpha_s\ge 2\\ \sum_i \alpha_i = k}} \frac{1}{\alpha_1!\cdots\alpha_s!}.$$

The number of ways of writing $k = \alpha_1 + \ldots + \alpha_s$ with each $\alpha_i \ge 2$ equals the number of ways of writing $k - s = \alpha_1' + \ldots + \alpha_s'$ where each $\alpha_i' \ge 1$ and is therefore $\binom{k-s}{s}$. Thus these remainder terms contribute

$$\leq \sum_{s<k/2} \frac{k!}{s!2^s}\binom{k-s}{s}\big(\log\log z + O(1)\big)^s. \tag{11}$$

Proposition 2 follows upon combining (9), (10), and (11).

The main novelty in our proof above is the introduction of the function $f_r(n)$ whose expectation over integers n below x is small unless r is squarefull. This leads easily to a recognition of the main term in the asymptotics of the moments. Previous approaches expanded out $(\omega(n) - \log\log x)^k$ using the binomial theorem, and then there are several main terms which must be carefully cancelled out before the desired asymptotic emerges. Our use of this simpler technique was inspired by (Montgomery and Soundararajan, 2004). Recently Rizwanur Khan (Khan, 2006) builds on this idea to prove that the spacings between normal numbers obey a Poisson distribution law.

This technique extends readily to the study of $\omega(n)$ in many other sequences. We formulate this in a sieve like setting:

Let $\mathcal{A} = \{a_1, \dots, a_x\}$ be a (multi)-set of x (not necessarily distinct) natural numbers. Let $\mathcal{A}_d = \#\{n \leq x : d \mid a_n\}$. We suppose that there is a real valued, non-negative multiplicative function $h(d)$ such that for square-free d we may write

$$\mathcal{A}_d = \frac{h(d)}{d}x + r_d.$$

It is natural to suppose that $0 \leq h(d) \leq d$ for all square-free d, and we do so below. Here r_d denotes a remainder term which we expect to be small: either small for all d, or maybe just small on average over d.

Let $\mathcal{P}$ be any set of primes. In sieve theory one attempts to estimate $\#\{n \leq x : (a_n, m) = 1\}$ for $m = \prod_{p\in\mathcal{P}} p$, in terms of the function h and the error terms r_d. Here we want to understand the distribution of values of $\omega_{\mathcal{P}}(a)$, as we vary through elements a of $\mathcal{A}$, where $\omega_{\mathcal{P}}(a)$ is defined to be the number of primes $p \in \mathcal{P}$ which divide a. We expect that the distribution of $\omega_{\mathcal{P}}(a)$ is normal with "mean" and "variance" given by

$$\mu_{\mathcal{P}} := \sum_{p\in\mathcal{P}} \frac{h(p)}{p} \quad \text{and} \quad \sigma_{\mathcal{P}}^2 := \sum_{p\in\mathcal{P}} \frac{h(p)}{p}\left(1 - \frac{h(p)}{p}\right),$$

and wish to find conditions under which this is true. There is a simple heuristic which explains why this should usually be true: Suppose that for each prime p we have a sequence of independent random variables $b_{1,p}, \dots, b_{x,p}$ each of which is 1 with probability $h(p)/p$ and 0 otherwise; and we let b_j be the product of the primes p for which $b_{j,p} = 1$. The b_j form a probabilistic model for the a_j satisfying our sieve hypotheses, the key point being that, in the model, whether or not b_j is divisible by different primes is independent.

One can use the central limit theorem to show that, as $x \to \infty$, the distribution of $\omega_{\mathcal{P}}(b)$ becomes normal with mean $\mu_{\mathcal{P}}$ and variance $\sigma_{\mathcal{P}}^2$.

PROPOSITION 3. *Uniformly for all natural numbers $k \le \sigma_{\mathcal{P}}^{2/3}$ we have*

$$\sum_{a\in\mathcal{A}} (\omega_{\mathcal{P}}(a)-\mu_{\mathcal{P}})^k = C_k x\sigma_{\mathcal{P}}^k\left(1+O\left(\frac{k^3}{\sigma_{\mathcal{P}}^2}\right)\right)+O\left(\mu_{\mathcal{P}}^k\sum_{d\in D_k(\mathcal{P})}|r_d|\right),$$

if k is even, and

$$\sum_{a\in\mathcal{A}} (\omega_{\mathcal{P}}(a)-\mu_{\mathcal{P}})^k \ll C_k x\sigma_{\mathcal{P}}^k\,\frac{k^{\frac{3}{2}}}{\sigma_{\mathcal{P}}}+\mu_{\mathcal{P}}^k\sum_{d\in D_k(\mathcal{P})}|r_d|,$$

if k is odd. Here $D_k(\mathcal{P})$ denotes the set of squarefree integers which are the product of at most k primes all from the set $\mathcal{P}$.

Proof. The proof is similar to that of Proposition 2, and so we record only the main points. We define $f_p(a) = 1 - h(p)/p$ if $p \mid a$ and $-h(p)/p$ if $p \nmid a$. If $r = \prod_i p_i^{\alpha_i}$ is the prime factorization of r we put $f_r(a) = \prod_i f_{p_i}(a)^{\alpha_i}$. Note that $\omega_{\mathcal{P}}(a) - \mu_{\mathcal{P}} = \sum_{p\in\mathcal{P}} f_p(a)$, and so

$$\sum_{a\in\mathcal{A}} (\omega_{\mathcal{P}}(a)-\mu_{\mathcal{P}})^k = \sum_{p_1,\ldots,p_k\in\mathcal{P}}\ \sum_{a\in\mathcal{A}} f_{p_1\cdots p_k}(a). \tag{12}$$

As in Proposition 2, consider more generally $\sum_{a\in\mathcal{A}} f_r(a)$. Suppose $r = \prod_{i=1}^s q_i^{\alpha_i}$ where the q_i are distinct primes and each $\alpha_i \ge 1$. Set $R = \prod_{i=1}^s q_i$ and observe that if $d = (a,R)$ then $f_r(a) = f_r(d)$. Note that

$$\begin{aligned}\sum_{\substack{a\in\mathcal{A}\\(a,R)=d}} 1 &= \sum_{a\in\mathcal{A}}\ \sum_{\substack{e|(R/d)\\ de|n}} \mu(e) = \sum_{e|R/d} \mu(e)\mathcal{A}_{de}\\ &= x\,\frac{h(d)}{d}\prod_{p|(R/d)}\left(1-\frac{h(p)}{p}\right)+\sum_{e|(R/d)}\mu(e)r_{de}.\end{aligned}$$

Therefore

$$\begin{aligned}\sum_{a\in\mathcal{A}} f_r(a) &= \sum_{d|R} f_r(d) \sum_{\substack{a\in\mathcal{A}\\(a,R)=d}} 1\\ &= x\sum_{d|R} f_r(d)\,\frac{h(d)}{d}\prod_{p|(R/d)}\left(1-\frac{h(p)}{p}\right)+\sum_{d|R} f_r(d)\sum_{e|(R/d)}\mu(e)r_{de}\\ &= G(r)x+\sum_{m|R} r_m E(r,m), \qquad (13)\end{aligned}$$

where

$$G(r) = \prod_{q^\alpha \| r} \left(\frac{h(q)}{q}\left(1 - \frac{h(q)}{q}\right)^{\alpha} + \left(\frac{-h(q)}{q}\right)^{\alpha}\left(1 - \frac{h(q)}{q}\right)\right), \tag{14}$$

and

$$E(r, m) = \prod_{\substack{q^\alpha \| r \\ q | m}} \left(\left(1 - \frac{h(q)}{q}\right)^{\alpha} - \left(\frac{-h(q)}{q}\right)^{\alpha}\right) \prod_{\substack{q^\alpha \| r \\ q | (R/m)}} \left(\frac{-h(q)}{q}\right)^{\alpha}. \tag{15}$$

We input the above analysis in (12). Consider first the main terms that arise. Notice that $G(r) = 0$ unless r is square-full, and so the main terms look exactly like the corresponding main terms in Proposition 2. We record the only small difference from the analysis there. When k is even there is a leading contribution from the terms with $s = k/2$ and all $\alpha_i = 2$ (in notation analogous to Proposition 2); this term contributes

$$\frac{k!}{2^{k/2}(k/2)!} \sum_{\substack{q_1, \dots, q_{k/2} \in \mathcal{P} \\ q_i \text{ distinct}}} \prod_{i=1}^{k/2} \frac{h(q_i)}{q_i}\left(1 - \frac{h(q_i)}{q_i}\right).$$

The sum over q's is bounded above by $\sigma_{\mathcal{P}}^k$, and is bounded below by

$$\left(\sum_{\substack{p \in \mathcal{P} \\ p \geq \pi_{k/2}(\mathcal{P})}} \frac{h(p)}{p}\left(1 - \frac{h(p)}{p}\right)\right)^{k/2} \geq (\sigma_{\mathcal{P}}^2 - k/8)^{k/2},$$

where we let $\pi_n(\mathcal{P})$ denote the n-th smallest prime in $\mathcal{P}$ and made use of the fact that $0 \leq (h(p)/p)(1 - h(p)/p) \leq 1/4$. The remainder of the argument is exactly the same as in Proposition 2.

Finally we need to deal with the "error" term contribution to (12). To estimate the error terms that arise in (12), we use that $|E(p_1 \cdots p_k, m)| \leq \prod_{p_i \nmid m} h(p_i)/p_i$. Thus the error term is

$$\leq \sum_{\ell=1}^{k} \sum_{\substack{m = q_1 \dots q_\ell \geq 1 \\ q_1 < q_2 < \dots < q_\ell \in \mathcal{P}}} |r_m| \sum_{\substack{p_1, \dots, p_k \in \mathcal{P} \\ m | p_1 \cdots p_k}} \prod_{p_i \nmid m} \frac{h(p_i)}{p_i}.$$

Fix m and let $e_j = \#\{i : p_i = q_j\}$ for each j, $1 \leq j \leq \ell$. Then there are $e_0 := k - (e_1 + \cdots + e_\ell) \leq k - \ell$ primes p_i which are not equal to any q_j, and so their contribution to the final sum is $\leq \mu_{\mathcal{P}}^{e_0}$. Therefore the final sum is

$$\le \sum_{0\le e_0\le k-\ell} \binom{k}{e_0} \mu_{\mathcal{P}}^{e_0} \sum_{\substack{e_1+\cdots+e_\ell=k-e_0\\ \text{each } e_i\ge 1}} \frac{(k-e_0)!}{e_1!\cdots e_\ell!}$$

$$\le \sum_{0\le e_0\le k-1} \binom{k}{e_0} \mu_{\mathcal{P}}^{e_0} \ell^{k-e_0} \le (\mu_{\mathcal{P}}+\ell)^k \ll 2\mu_{\mathcal{P}}^k,$$

since $k^3 \le \sigma_{\mathcal{P}}^2 \le \mu_{\mathcal{P}}$. This completes the proof of the proposition.

One way of using Proposition 3 is to take $\mathcal{P}$ to be the set of primes below z where z is suitably small so that the error term arising from the $|r_d|$'s is negligible. If the numbers a in $\mathcal{A}$ are not too large, then there cannot be too many primes larger than z that divide a, and so Proposition 3 furnishes information about $\omega(a)$. Note that we used precisely such an argument in deducing Theorem 1 from Proposition 2.

In this manner, Proposition 3 may be used to prove the Erdős-Kac theorem for many interesting sequences of integers. For example, Halberstam (Halberstam, 1956) showed such a result for the shifted primes $p-1$, which the reader can now deduce from Proposition 3 and the Bombieri–Vinogradov theorem.

Similarly, one can take $\mathcal{A} = \{f(n) : n \le x\}$ for $f(t) \in \mathbb{Z}[t]$. In this case $h(p)$ is bounded by the degree of f except at finitely many primes, and the prime ideal theorem implies that $\mu_{\mathcal{P}}, \sigma_{\mathcal{P}} = m \log\log x + O(1)$ where m is the number of distinct irreducible factors of f. Again this example was first considered by Halberstam (Halberstam, 1956).

Alladi (Alladi, 1987) proved an Erdős–Kac theorem for integers without large prime factors. Proposition 3 reduces this problem to obtaining information about multiples of d in this set of “smooth numbers.” We invite the reader to fill in this information.

In place of $\omega(a)$ we may study more generally the distribution of values of $g(a)$ where g is an “additive function.” Recall that an additive function satisfies $g(1) = 0$, and $g(mn) = g(m) + g(n)$ whenever m and n are coprime. Its values are determined by the prime-power values $g(p^k)$. If in addition $g(p^k) = g(p)$ for all $k \ge 1$ we say that the function g is “strongly additive.” The strongly additive functions form a particularly nice subclass of additive functions and for convenience we restrict ourselves to this subclass.

PROPOSITION 4. *Let $\mathcal{A}$ be a (multi)-set of x integers, and let $h(d)$ and r_d be as above. Let $\mathcal{P}$ be a set of primes, and let g be a real-valued, strongly* *additive function with $|g(p)| \le M$ for all $p \in \mathcal{P}$. Let*

$$\mu_{\mathcal{P}}(g) = \sum_{p\in\mathcal{P}} g(p)\frac{h(p)}{p}, \quad \textit{and} \quad \sigma_{\mathcal{P}}(g)^2 = \sum_{p\in\mathcal{P}} g(p)^2\frac{h(p)}{p}\Big(1 - \frac{h(p)}{p}\Big).$$

Then, uniformly for all even natural numbers $k \le (\sigma_{\mathcal{P}}(g)/M)^{2/3}$,

$$\sum_{a\in\mathcal{A}}\Big(\sum_{\substack{p|a\\ p\in\mathcal{P}}} g(p) - \mu_{\mathcal{P}}(g)\Big)^k = C_k x\sigma_{\mathcal{P}}(g)^k\Big(1 + O\Big(\frac{k^3M^2}{\sigma_{\mathcal{P}}(g)^2}\Big)\Big) + O\Big(M^k\Big(\sum_{p\in\mathcal{P}}\frac{h(p)}{p}\Big)^k \sum_{d\in D_k(\mathcal{P})}|r_d|\Big),$$

while for all odd natural numbers $k \le (\sigma_{\mathcal{P}}(g)/M)^{2/3}$,

$$\sum_{a\in\mathcal{A}}\Big(\sum_{\substack{p|a\\ p\in\mathcal{P}}} g(p) - \mu_{\mathcal{P}}(g)\Big)^k \ll C_k x\sigma_{\mathcal{P}}(g)^k\frac{k^{3/2}M}{\sigma_{\mathcal{P}}(g)} + M^k\Big(\sum_{p\in\mathcal{P}}\frac{h(p)}{p}\Big)^k \sum_{d\in D_k(\mathcal{P})}|r_d|.$$

Proof. We follow closely the proofs of Propositions 2 and 3, making appropriate modifications. Let $f_r(n)$ be as in the proof of Proposition 3. Then we wish to evaluate

$$\sum_{a\in\mathcal{A}}\Big(\sum_{p\in\mathcal{P}} g(p)f_p(a)\Big)^k = \sum_{p_1,\dots,p_k\in\mathcal{P}} g(p_1)\cdots g(p_k)\sum_{a\in\mathcal{A}} f_{p_1\cdots p_k}(a).$$

We may now input the results (13, 14, 15) here. Consider first the error terms that arise. Since $|g(p)| \le M$ for all $p \in \mathcal{P}$ this contribution is at most M^k times the corresponding error in Proposition 5. To wit, the error terms are

$$\ll M^k\Big(\sum_{p\in\mathcal{P}}\frac{h(p)}{p}\Big)^k \sum_{d\in D_k(\mathcal{P})}|r_d|.$$

As for the main term, note that $G(r) = 0$ unless r is square-full and so if $q_1 < q_2 < \ldots < q_s$ are the distinct primes among the $p_1, \ldots, p_k$ our main term is

$$x\sum_{s\le k/2}\sum_{\substack{q_1<\ldots<q_s\\ q_i\in\mathcal{P}}}\sum_{\substack{\alpha_1,\dots,\alpha_s\ge 2\\ \sum_i\alpha_i=k}}\frac{k!}{\alpha_1!\cdots\alpha_s!}\prod_{i=1}^{s} g(q_i)^{\alpha_i}G(q_1^{\alpha_1}\cdots q_s^{\alpha_s}). \tag{16}$$

When k is even there is a term with $s = k/2$ and all $\alpha_i = 2$ which is the leading contribution to (16). This term contributes

$$x\frac{k!}{2^{k/2}(k/2)!}\sum_{\substack{q_1,\dots,q_{k/2}\in\mathcal{P}\\ q_i \text{ distinct}}}\prod_{i=1}^{k/2} g(q_i)^2\frac{h(q_i)}{q_i}\Big(1 - \frac{h(q_i)}{q_i}\Big).$$

If we fix $q_1, \ldots, q_{k/2-1}$, then the sum over $q_{k/2}$ is $\sigma_{\mathcal{P}}(g)^2 + O(M^2 k)$, since $|g(p)| \le M$ for all $p \in \mathcal{P}$, and $0 \le h(p) \le p$. Therefore the contribution of the term $s = k/2$ to (16) is

$$C_k x(\sigma_{\mathcal{P}}(g)^2 + O(M^2 k))^{k/2} = C_k x \sigma_{\mathcal{P}}(g)^k \left(1 + O\left(\frac{M^2 k^2}{\sigma_{\mathcal{P}}(g)^2}\right)\right),$$

since $kM \le \sigma_{\mathcal{P}}(g)$.

Now consider the terms $s < k/2$ in (16). Since $|G(q_1^{\alpha_1} \cdots q_s^{\alpha_s})| \le \prod_{i=1}^{s} (h(q_i)/q_i)(1 - h(q_i)/q_i)$, and $\prod_{i=1}^{s} |g(q_i)|^{\alpha_i} \le M^{k-2s} \prod_{i=1}^{s} |g(q_i)|^2$, we see that these terms contribute an amount whose magnitude is

$$\le x \sum_{s<k/2} \frac{k!}{s!} M^{k-2s} \left(\sum_{q \in \mathcal{P}} |g(q)|^2 \frac{h(q)}{q} \left(1 - \frac{h(q)}{q}\right)\right)^s \sum_{\substack{\alpha_1, \ldots, \alpha_s \ge 2 \\ \sum \alpha_i = k}} \frac{1}{\alpha_1! \cdots \alpha_s!}$$

$$\le x \sum_{s<k/2} \frac{k!}{s! 2^s} \binom{k-s}{s} M^{k-2s} \sigma_{\mathcal{P}}(g)^{2s},$$

using that $\binom{k-s}{s}$ equals the number of ways of writing $k = \sum \alpha_i$ with each $\alpha_i \ge 2$. The proposition follows.

One way to apply Proposition 4 is to take $\mathcal{P}$ to be the set of all primes below z with $|g(p)|$ small. If there are not too many values of p with $|g(p)|$ large, then we would expect that $g(a)$ is roughly the same as $g_{\mathcal{P}}(a)$ for most a. In such situations, Proposition 4 which furnishes the distribution of $g_{\mathcal{P}}(a)$ would also furnish the distribution of $g(a)$. In this manner one can deduce the result of Kubilius and Shapiro (Shapiro, 1956) which is a powerful generalization of the Erdős–Kac theorem for additive functions. Indeed we can derive such a Kubilius–Shapiro result in the more general sieve theoretic framework given above, and for all additive functions rather than only for the subclass of strongly additive functions.

There are many other interesting number theory questions in which an Erdős–Kac type theorem has been proved. We have collected some of these references below[1] and invite the reader to determine which of these Erdős–Kac type theorems can be deduced from the results given herein. The reader may also be interested in the textbooks (Elliott, 1979; Kubilius, 1964; Tenenbaum, 1995) for a more classical discussion of some of these issues, and to the elegant essays (Billingsley, 1973; Kac, 1959).

[1] Thanks are due to Yu-Ru Liu for her help with this.

Acknowledgements

Le premier auteur est partiellement soutenu par une bourse du Conseil de recherches en sciences naturelles et en génie du Canada. The second author is partially supported by the National Science Foundation.

References

Alladi, K. (1987) An Erdős–Kac theorem for integers without large prime factors, *Acta Arith.* **49**, 81–105.

Billingsley, P. (1973) Prime numbers and Brownian motion, *Amer. Math. Monthly* **80**, 1099–1115.

David, C. and Pappalardi, F. (1999) Average Frobenius distributions of elliptic curves, *Internat. Math. Res. Notices* **1999**, 165–183.

Elliott, P. D. T. A. (1979) *Probabilistic number theory.* Vol. I. and II, Vol. 239 and 240 of *Grundlehren Math. Wiss.*, New York–Berlin, Springer.

Elliott, P. D. T. A. and Sárkőzy, A. (1997) The distribution of the number of prime divisors of numbers of form $ab + 1$, In *New trends in probability and statistics.* Vol. 4, Palanga, 1996, pp. 313–321, VSP, Utrecht.

Erdős, P. (1935) On the normal order of prime factors of $p - 1$ and some related problems concerning Euler's φ-functions, *Quart. J. Math.(Oxford)* **6**, 205–213.

Erdős, P. and Kac, M. (1940) The Gaussian law of errors in the theory of additive number theoretic functions, *Amer. J. Math* **62**, 738–742.

Erdős, P., Maier, H., and Sárkőzy, A. (1987) On the distribution of the number of prime factors of sums $a + b$, *Trans. Amer. Math. Soc* **302**, 269–280.

Erdős, P. and Pomerance, C. (1985) On the normal number of prime factors of $\varphi(n)$, *Rocky Mountain J. Math* **15**, 343–352.

Erdős, P. and Wintner, A. (1939) Additive arithmetical functions and statistical independence, *Amer. J. Math.* **61**, 713–721.

Halberstam, H. (1955) On the distribution of additive number theoretic functions. I, *J. London Math. Soc.* **30**, 43–53.

Halberstam, H. (1956) On the distribution of additive number theoretic functions. III, *J. London Math. Soc.* **31**, 15–27.

Hardy, G. H. and Ramanujan, S. (1917) The normal number of prime factors of a number n, *Quar. J. Pure. Appl. Math* **48**, 76–97.

Hensley, D. (1994) The number of steps in the Euclidean algorithm, *J. Number Theory* **49**, 142–182.

Hildebrand, A. (1987) On the number of prime factors of integers without large prime divisors, *J. Number Theory* **25**, 81–106.

Kac, M. (1959) *Statistical independence in probability, analysis and number theory*, Vol. 12 of *Carus Math. Monogr.*, New York, Math. Assoc. America.

Khan, R. (2006) On the distribution of normal numbers, preprint.

Kubilius, J. (1964) *Probabilistic methods in the theory of numbers*, Vol. 11 of *Transl. Math. Monogr.*, Providence, RI, Amer. Math. Soc.

Kuo, W. and Liu, Y.-R. (2006) Erdős–Pomerance's conjecture on the Carlitz module, to appear.

Li, S. and Pomerance, C. (2003) On generalizing Artin's conjecture on primitive roots to composite moduli, *J. Reine Angew. Math.* **556**, 205–224.

Liu, Y.-R. (2004) A generalization of the Erdős–Kac theorem and its applications, *Canad. Math. Bull.* **47**, 589–606.

Liu, Y.-R. (2005a) A prime analogue of Erdős–Pomerance's conjecture for elliptic curves, *Comment. Math. Helv.* **80**, 755–769.

Liu, Y.-R. (2005b) Prime divisors of the number of rational points on elliptic curves with complex multiplication, *Bull. London Math. Soc* **37**, 658–664.

Mauduit, C. and Sárkőzy, A. (1996) On the arithmetic structure of sets characterized by sum of digits properties, *J. Number Theory* **61**, 25–38.

Montgomery, H. and Soundararajan, K. (2004) Primes in short intervals, *Comm. Math. Phys.* **252**, 589–617.

Murty, M. R. and Saidak, F. (2004) Non-abelian generalizations of the Erdős–Kac theorem, *Canad. J. Math* **56**, 356–372.

Murty, V. K. and Murty, M. R. (1984a) An analogue of the Erdős–Kac theorem for Fourier coefficients of modular forms, *Indian J. Pure Appl. Math.* **15**, 1090–1101.

Murty, V. K. and Murty, M. R. (1984b) Prime divisors of Fourier coefficients of modular forms, *Duke Math. J.* **51**, 57–76.

Sathe, L. G. (1953) On a problem of Hardy on the distribution of integers having a given number of prime factors. II., *J. Indian Math. Soc. (N.S.)* **17**, 83–141.

Selberg, A. (1954) Note on a paper by L. G. Sathe, *J. Indian Math. Soc. (N.S.)* **18**, 83–87.

Shapiro, H. (1956) Distribution functions of additive arithmetic functions, *Proc. Nat. Acad. Sci. USA* **42**, 426–430.

Tenenbaum, G. (1995) *Introduction to analytic and probabilistic number theory*, Vol. 46 of *Cambridge Stud. Adv. Math.*, Cambridge, Cambridge University Press.

Turán, P. (1934) On a theorem of Hardy and Ramanujan, *J. London Math. Soc.* **9**, 274–276.

UNIFORM DISTRIBUTION, EXPONENTIAL SUMS, AND CRYPTOGRAPHY

John B. Friedlander
University of Toronto

In these notes we discuss various sequences of numbers which are motivated by cryptographic considerations. This suggests the study of their uniform distribution and, in turn, the bounding of relevant exponential sums. Several of the bounds we give have since been quantitatively sharpened, by Garaev (Garaev, 2005) and, spectacularly so, in recent work of Bourgain (Bourgain, 2004; Bourgain, 2005).

1. Randomness and Pseudorandomness

A random sequence, for example of numbers, is one which carries with it the notion of unpredictability; the basic idea is that it should not be possible to guess, at any given stage, from the history of the sequence, what will be the future of the sequence. Random sequences occur, for example, from the repeated tosses of a fair coin or a fair die.

For various reasons, such as a subroutine in some cryptographic protocol, one would like to have a supply of random sequences at one's fingertips. Thanks to the speed of modern computers, large numbers need to be involved in such applications and, since it is not practical to sit around repeatedly flipping a coin, one uses instead machine-generated sequences, hence deterministic, yet having the "appearance of randomness." We call such a sequence "pseudorandom." We expect such a sequence to be finite, in other words periodic, in which case we want it to have a large period to preserve the illusion of randomness. We would also like some other measures of randomness. For the most part, we expect necessary criteria. The problem of finding quantifiable measures guaranteeing a particular sequence with a finite period behaves in a truly random fashion seems much more difficult.

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 29–57.

2. Uniform Distribution and Exponential Sums

The simplest criterion one might think of is uniform distribution. Let, for instance, u_n be a sequence of points in $[0, 1)$.

DEFINITION 2.1. We say that u_n is uniformly distributed if for every $0 \leq a < b < 1$ we have

$$\sum_{\substack{n \leq x \\ a < u_n \leq b}} 1 \sim (b - a)x.$$

Certainly the fact that a sequence of numbers is uniformly distributed in this fashion does not at all imply randomness. There are many very simple deterministic sequences which are uniformly distributed. Still, one would think that this is the least we could ask of a random sequence; the fact that it bunched up in certain locations would make it to at least some extent predictable.

Following the pioneering work of Hermann Weyl, the notion of uniform distribution leads us inexorably to the topic of exponential sums. This is because of the following theorem and its many generalizations:

2.1. THE WEYL CRITERION

The sequence u_n is uniformly distributed if and only if, for every integer $h \neq 0$, we have

$$\sum_{n \leq x} e(hu_n) = o(x).$$

Here we have used the standard notation

$$e(u) = e^{2\pi iu},$$

which will occur again and again.

We next want to both generalize and quantify the notion of uniform distribution. Consider a finite sequence of N points in the n-dimensional unit cube $[0, 1)^n$:

$$\Gamma = (\gamma_{0j}, \gamma_{1j}, \ldots, \gamma_{(n-1)j}), \quad j = 1, 2, \ldots, N,$$

still not the most general situation, but sufficient for our current purposes.

We define the “discrepancy” of the sequence Γ as:

$$\Delta_\Gamma = \sup_{B \subseteq [0,1)^n} \left| \frac{\mathcal{N}_\Gamma(B)}{N} - |B| \right|$$

where $\mathcal{N}_\Gamma(B)$ denotes the number of points in $\Gamma \cap B$, where $|B|$ is the volume of B, and where the supremum runs over all boxes of type

$$B = [\alpha_0, \beta_0] \times \cdots \times [\alpha_{n-1}, \beta_{n-1}) \subset [0, 1)^n.$$

The obvious trivial bound for this discrepancy is $\Delta_\Gamma \leq 1$. A small discrepancy, that is

$$\Delta_\Gamma = o(1),$$

is equivalent to uniform distribution. We study this, again following H. Weyl, via exponential sums.

Let $\underline{a} = (a_0, \ldots, a_{n-1}) \in \mathbb{Z}^n$. Here, the relevant exponential sum is

$$S_{\underline{a}}(\Gamma) = \sum_{j=1}^{N} e\left(\sum_{i=0}^{n-1} a_j \gamma_{ij}\right).$$

In this situation the Weyl criterion becomes:

THEOREM 2.2. *The sequence Γ is uniformly distributed if and only if, for every vector $\underline{a} \neq \underline{0}$, we have*

$$\lim_{N \to \infty} N^{-1} S_{\underline{a}}(\Gamma) = 0.$$

Still smaller discrepancies, say $\Delta_\Gamma \ll |B|^{-\alpha}$ for some fixed $\alpha > 0$, provide a quantitative measure of just how uniform the distribution is. This more general notion of discrepancy is also approached through exponential sums. A basic bound for the discrepancy, in terms of the exponential sums $S_{\underline{a}}(\Gamma)$, is given by the following result.

LEMMA 2.3. *There exists an absolute constant $C > 0$ such that, for all $L \in \mathbb{N}$ we have*

$$\Delta_\Gamma < C^n \left(\frac{1}{L} + \frac{1}{N} \sum_{\substack{\underline{a} \neq 0 \\ \max_i |\underline{a}_i| \leq L}} \prod_{i=0}^{n-1} (1 + |a_i|)^{-1} |S_{\underline{a}}(\Gamma)| \right).$$

In the one-dimensional case this is the well-known Erdös–Turan inequality.

3. Exponential Sums and Cryptography

In this section we give examples of some exponential sums which arise in connection with two of the most basic cryptographic considerations.

(A) The discrete logarithm problem.

Suppose we are given a prime p and a primitive root g, that is a generator of the (cyclic) unit group $\mathcal{U}_p = \mathbb{Z}_p^\times$. The exponential map $x \to g^x \pmod p$ is computationally easy, by the process of repeated squaring. The inverse map, given g^x recover x, is the 'discrete logarithm' map and the discrete logarithm problem is the still unsolved question of deciding whether this can be computed quickly, say in polynomial time, that is bounded by $(\log p)^N$ for some fixed N. This problem suggests that we study the ordered pairs of fractional parts $(\{x/p\}, \{g^x/p\})$, $x = 1, \ldots, p-1$, where this normalization of the data places our points in the unit square $[0, 1)^2$.

It has long been known that this sequence of pairs is uniformly distributed. In fact, we don't need to use a primitive root g; it is sufficient to take an element θ having a large multiplicative order, say t. As a result we can ask the same question, not just for prime numbers p, but also for arbitrary integer. Although these typically won't have any primitive roots they will usually have an element of large order.

In studying these pairs the relevant exponential sum is

$$\sum_{x=1}^{t} e_p(b\theta^x)e_t(ax),$$

where, having earlier defined $e(u) = \exp(2\pi iu)$, we now require the additional notation $e_m(u) = e(u/m)$.

These exponential sums have been studied by many people, for example Korobov, Konyagin, Heath–Brown, Bourgain. The first published application of this sum to the discrete logarithm seems to be due to Shparlinski (Shparlinski, 2002) who showed that $\Delta = o(1)$ under the condition $t > p^{1/3+\varepsilon}$, since relaxed, most recently by the work of Bourgain, to $t > p^{\varepsilon}$.

(B) The Diffie–Hellman key exchange.

Considered by many the foundation stone of public-key cryptography, this is a procedure by which two parties, Alice (A) and Bob (B), share a secret key which they can set up while communicating over an insecure line. They begin by agreeing (in public) on a large prime p and a primitive root g modulo p.

Then, to set up the key:

- A chooses x at random, and sends (the least positive residue of) g^x to B,
- B chooses y at random and sends g^y to A,

- A, knowing x and g^y computes g^{yx} modulo p,
- B, knowing y and g^x computes $g^{xy} = g^{yx}$ modulo p.

Here the point is that any third party, even knowing p, g, g^x, g^y, hopefully cannot compute g^{xy} in a reasonable time. One doesn't know whether this is the case in general. It is easy to see that, if the discrete logarithm problem can be quickly solved, then so can this one. It might possibly be the case that the reverse is true as well or, alternatively it might be that this one can be broken without breaking the discrete logarithm, but this too is unknown.

Actually what is at issue in such a situation is not so much the number itself but rather its (binary) representation and one would like to know that the interloper can't compute many of these bits, not just that he can't compute all of them. Alternatively, one would like to know that he can't compute these bits with high probability.

A desirable feature for this is that one cannot distinguish the triple (g^x, g^y, g^{xy}) from a random triple in $(\mathbb{Z}_p^\times)^3$, at least in the sense that these triples are uniformly distributed. One may remark that actually these triples are not completely random, for example, if g^x is a quadratic residue modulo p then certainly so is g^{xy}. Hence, for technical reasons it is useful to replace g by a residue class θ of high multiplicative order which is an rth power for all small $r \mid p-1$. Hence, we certainly want $p-1$ to have some large prime factor. Once we have done this, as with the discrete logarithm, the set-up now makes sense even for general modulus m. We shall return to this later but for the time being we consider prime modulus.

Denote $t = \operatorname{ord}\theta$ and consider the normalized Diffie–Hellman triples

$$\left(\frac{\theta^x}{p}, \frac{\theta^y}{p}, \frac{\theta^{xy}}{p}\right)$$

$x = 1, 2, \ldots, t$, $y = 1, 2, \ldots, t$.

The Weyl criterion now reads as follows: The triples above are uniformly distributed in the unit cube if and only if, for all non-zero triples $(a, b, c) \neq (0, 0, 0)$, we have

$$S_{abc}(p, t) = o(t^2)$$

where S is the exponential sum

$$S_{abc}(m, t) \doteq \sum_{x=1}^{t} \sum_{y=1}^{t} e_m(a\theta^x + b\theta^y + c\theta^{xy}).$$

4. Some Exponential Sum Bounds

Actually, we shall bound the sum

$$V_{ac}(p,t) \doteq \sum_{y=1}^{t} \left| \sum_{x=1}^{t} e_p(a\theta^x + c\theta^{xy}) \right|^4 .$$

A bound for this is given by the following result; see (Canetti et al., 1999; Canetti et al., 2000) for the results of this section.

THEOREM 4.1. *For $(a,c) \neq (0,0)$ we have*

$$V_{ac}(p,t) \ll t^{11/3} p.$$

Note that the trivial bound is t^5 so the above result provides some saving in case that $t > p^{3/4+\varepsilon}$.

The proof of Theorem 4.1 is more lengthy than the other results presented here and we postpone our sketch of the ideas involved to the final section of the paper.

From this result we obtain a uniformity of distribution statement for the Diffie–Hellman triples. Specifically,

THEOREM 4.2. *If $t > p^{3/4+\varepsilon}$ the triples*

$$\left(\frac{\theta^x}{p}, \frac{\theta^y}{p}, \frac{\theta^{xy}}{p} \right)$$

$x = 1, \ldots, t$, $y = 1, \ldots, t$ are uniformly distributed in the unit cube.

It is easy to see that Theorem 4.1 implies Theorem 4.2 by means of Hölder's inequality. Indeed,

$$\begin{aligned} |S_{abc}(p,t)| &= \left| \sum_{x=1}^{t} \sum_{y=1}^{t} e_p(a\theta^x + b\theta^y + c\theta^{xy}) \right| \\ &\leq \sum_{y=1}^{t} \left| \sum_{x=1}^{t} e_p(a\theta^x + c\theta^{xy}) \right| \\ &\leq t^{3/4} (V_{ac}(p,t))^{1/4} \ll t^{5/3} p^{1/4}, \end{aligned}$$

so that the Weyl criterion implies the result.

We remark that, more recently, Theorem 4.2 has been improved by Garaev who replaced $\frac{3}{4} + \varepsilon$ by $\frac{1}{2} + \varepsilon$ and by Bourgain who reduced it to ε.

One learns a little more by showing uniformity of distribution results for various subsets of the triples. For example, let $\underline{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ be three binary strings of length k_1, k_2, k_3 and let

$$L_t(\underline{\sigma}) = \#\{(x, y), x = 1, \dots, t, y = 1, \dots, t\}$$

count the number of pairs for which σ_1, σ_2, σ_3 is the string of k_1, k_2, k_3 least significant bits of θ^x, θ^y, θ^{xy} respectively.

THEOREM 4.3. *We have*

$$|L_t(\underline{\sigma}) - t^2 2^{-k_1-k_2-k_3}| \ll t^{5/3} p^{1/4} (\log p)^3.$$

Hence, provided that $t > p^{3/4+\delta}$, for a positive proportion (depending on δ), the least significant bits of the Diffie–Hellman triples are uniformly distributed. A similar result holds if we consider instead the 'most' significant bits, and a somewhat weaker result for subsets of bits in general position.

5. General Modulus and Discrepancy of Diffie–Hellman Triples

In work with Konyagin and Shparlinski (Friedlander et al., 2002) we considered the problem of extending the above results on Diffie–Hellman triples to composite moduli (which might be useful for example in the case of the product of two large primes) and also asking for bounds on the discrepancy rather than simply requesting uniform distribution. Let $m \in \mathbb{N}$, $\theta \in \mathcal{U}_m$ $(= \mathbb{Z}_m^\times)$ and, as before, let t be the multiplicative order of θ modulo m. Denote by D_t the discrepancy of the set of triples

$$\left(\left\{\frac{\theta^x}{m}\right\}, \left\{\frac{\theta^y}{m}\right\}, \left\{\frac{\theta^{xy}}{m}\right\}\right)$$

of fractional parts as $x, y = 1, 2, \dots, t$. We have the following result.

THEOREM 5.1. *Let $\varepsilon > 0$. Then*

$$D_t \underset{\varepsilon}{\ll} t^{-11/16} m^{5/8+\varepsilon},$$

where, as indicated, the implied constant may depend on ε.

Since the bound $D_t \leq 1$ is trivial, the above theorem is non-trivial if $t > m^{10/11+\varepsilon}$. In case $m = p$ is prime, recall we had a non-trivial bound in a much wider range.

By Lemma 2.3 we reduce this problem to an exponential sum, specifically

$$D_t \ll m^{-1} + t^{-2} \sum_{\substack{-m<a,b,c<m \\ (a,b,c)\neq(0,0,0)}} \frac{|S_{abc}|}{(1+|a|)(1+|b|)(1+|c|)},$$

where $S_{abc} = S_{abc}(m,t)$ is as before.

For this latter sum we can prove:

THEOREM 5.2. *If the greatest common divisor* $(ac,m) = 1$, *then*

$$S_{abc}(m,t) \ll_{\varepsilon} t^{21/16} m^{5/8+\varepsilon}.$$

Actually, proceeding along the lines of the prime modulus case, we deduce this from a bound for the sum

$$W_{ac}(m,t) = \sum_{y=1}^{t} \left| \sum_{x=1}^{t} e_m(a\theta^x + c\theta^{xy}) \right|$$

and then use the triangle inequality in the form

$$|S_{abc}| \leq \min\{W_{ac}, W_{bc}\}.$$

Our bound for the sum W is again deduced via Hölder's inequality from a bound for the fourth moment, in this case the following result.

THEOREM 5.3.

$$V_{ac}(m,t) = \sum_{y=1}^{t} \left| \sum_{x=1}^{t} e_m(a\theta^x + c\theta^{xy}) \right|^4 \ll_{\varepsilon} (a,m) t^{9/4} m^{5/2+\varepsilon}.$$

Modifications of the method of proof of the above results lead to estimates for incomplete sums. (See (Banks et al., 2006b) for details in the case of prime modulus.) One obtains bounds for the sum

$$\sum_{y\in\mathcal{Y}} \left| \sum_{x\in\mathcal{X}} e_m(a\theta^x + c\theta^{xy}) \right|^4$$

over the intervals $\mathcal{X} = [L_0, L_0 + L]$, $\mathcal{Y} = [K_0, K_0 + K]$ and these in turn give bounds for the incomplete sums

$$\sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} e_m(a\theta^x + b\theta^y + c\theta^{xy}).$$

These are in most cases (that is, unless the intervals are very close to the full period), better than those which would be obtained simply by the standard method of "completing the sum" and then applying our earlier bounds for the complete sums S_{abc}.

6. Pseudorandom Number Generation

One source of sequences of numbers important for applications to cryptography is given by those which are spun off by various pseudorandom number generators. A well-known example of this is the RSA generator.

We begin with integers α the root, m the modulus, and e the exponent, each at least 2 to avoid trivialities, and satisfying the coprimality conditions $(\alpha, m) = 1$ and $(e, \varphi(m)) = 1$. (The notation e for the exponent is very standard. Hopefully it will not cause confusion with the exponential function which, throughout, will always be occurring with a subscript. As usual φ is Euler's function.)

We define the sequence u_n, $n = 0, 1, 2, \ldots$, by beginning with the root $u_0 = \alpha$ and then successively exponentiating:

$$u_n \equiv u_{n-1}^e \pmod{m}, \quad 0 \le u_n < m,$$

so that

$$u_n \equiv \alpha^{e^n} \pmod{m}.$$

It is not hard to see that this sequence is purely periodic, say with period t. We would like to know that the sequence u_n is uniformly distributed modulo m. Again we come back to the Weyl criterion. Now, the relevant exponential sum is

$$S_a = \sum_{n=1}^{t} e_m(au_n).$$

In practice we usually take m to be the product of two large primes. We have the following result from (Friedlander et al., 1999).

THEOREM 6.1. *Let m be a prime or the product of two primes. Let $(a, m) = 1$. Then*

$$S_a \ll m^{23/24}.$$

REMARK. If m is the product of k primes with $(k \ge 3)$ the method gives a weaker but still non-trivial bound (with an exponent $f(k) < 1$).

COROLLARY 6.2. *If m is prime or the product of two primes and $t > m^{23/24+\delta}$ then $\{u_n\}$ is uniformly distributed modulo m and a positive proportion of its least (most) significant bits is uniformly distributed.*

Because the proof is not very long, yet illustrates some of the basic ideas used in this area, we show how the above Theorem 6.1 follows from the exponential sum bound given earlier in Theorem 4.1.

We begin by noting that, since u_n has period t, it follows that for every $k \in \mathbb{N}$,

$$S_a \doteq \sum_{n=1}^{t} e_m(au_n) = \sum_{n=1}^{t} e_m(au_{n+k}) = \sum_{n=1}^{t} e_m(a\alpha^{e^{n+k}}) = \sum_{n=1}^{t} e_m(ax^{e^n}),$$

where $x = \alpha^{e^k}$. Since the α^{e^k} are distinct modulo m for $k = 1, \ldots, t$ we deduce that

$$\begin{aligned} t|S_a|^2 &= \sum_{k=1}^{t} \left| \sum_{n=1}^{t} e_m(au_{n+k}) \right|^2 \leq \sum_{x=0,(x,m)=1}^{m-1} \left| \sum_{n=1}^{t} e_m(ax^{e^n}) \right|^2 \\ &= \sum_{x=0,(x,m)=1}^{m-1} \sum_{n,k=1}^{t} e_m(a(x^{e^n} - x^{e^k})) \leq \sum_{n,k=1}^{t} \left| \sum_{x=0,(x,m)=1}^{m-1} e_m(a(x^{e^n} - x^{e^k})) \right| \\ &\leq \sum_{k=1}^{t} \sum_{y=1,(y,\varphi(m))=1}^{\varphi(m)} \left| \sum_{x=0,(x,m)=1}^{m-1} e_m(a(x^{ye^k} - x^{e^k})) \right|, \end{aligned}$$

writing $y = e^{n-k}$. Because we have the condition $(e, \varphi(m)) = 1$, it follows that the inner sum is independent of k. Hence, replacing x^{e^k} by z,

$$|S_a|^2 \leq \sum_{y=1}^{\varphi(m)} \left| \sum_{z=0,(z,m)=1}^{m-1} e_m(a(z^y - z)) \right|.$$

At this point we are going to simplify slightly and treat the case where m is prime (the other is only slightly more complicated, requiring little more other than the Chinese Remainder Theorem). By applying Hölder's inequality to the sum on the right hand side of the previous inequality, we obtain

$$|S_a|^2 \leq \varphi(m)^{3/4} \left\{ \sum_{y=1}^{\varphi(m)} \left| \sum_{z=0,(z,m)=1}^{m-1} e_m(a(z^y - z)) \right|^4 \right\}^{1/4}.$$

If we then make a change of variable replacing z by g^ζ we arrive for $m = p$, $t = p - 1$, at the special case $a = -c$ of the sum $V_{ac}(t)$ considered in Theorem 4.1.

The RSA generator is a special case of the "power generator" of pseudo-random numbers. Amongst the latter, the RSA generator is characterized by having the additional coprimality condition $(e, \varphi(m)) = 1$. As we just saw, this condition is crucial to the previous proof. Yet, there are many interesting cases of the power generator for which this condition does not hold. Indeed the first such generator considered, and certainly the most obvious choice, is the Blum–Blum–Shub generator, wherein one takes the exponent to be $e = 2$ so that this condition never holds in practice.

In later work with Shparlinski (Friedlander and Shparlinski, 2001) we were able to treat the general case. We obtained the following result.

THEOREM 6.3. *Let $\delta > 0$, let $(e, \varphi(m))$ be arbitrary and let (m be prime or) $m = p\ell$ be the product of two primes. Define $\Delta = (p-1, \ell-1)$. If $t/\Delta > m^{3/4+\delta}$, then the sequence u_n is uniformly distributed modulo m.*

Here the main tool in the proof is a very well-known exponential sum bound due to A. Weil. We recall that, if p is prime, $f \in \mathbb{Z}[x]$ with degree $\deg f = d$, and such that f is non-constant modulo p, then

$$\left| \sum_{x=1}^{p} e_p(f(x)) \right| < dp^{1/2}.$$

This bound is very good if the degree d is not large. In our case we have u_n is given by $u_0^{e^n}$ and this leads to polynomials having monomials x^{e^k} of very high degree. For these the Weil bound is worse than the trivial bound p. To get around this we use the following lemma, the proof of which follows quickly from the box principle and elementary sieving.

LEMMA 6.4. *Let $\delta > 0$, $(e, T) = 1$ and let τ denote the multiplicative order of e $(\bmod T)$. Then, for $h \geq T^{\delta}$, there exists r with $(r, T) = 1$ such that the congruence $re^j \equiv y \pmod T$, $1 \leq j \leq \tau$, $0 \leq y < h$ has $\gg \tau h/T$ solutions.*

The usefulness of the lemma is as follows. We want, as in the previous proof, to make translations of the variable so as to produce many copies of our exponential sum. However, if we then want to apply Weil's bound the result might usually be bad due to the size of the degree. Note however that, in the previous proof, there was no necessity to translate by every value of the variable k; we could have done this for a subset of these, obtaining however a weaker result. The existence of the integer r in the lemma tells us that there exists an initial change of variables, after which a large number of the translations will result in polynomials of degree not too large. We can no longer make as many copies of the sum as before, but we get to apply the stronger Weil bound to the copies we have made.

As in the case of the Diffie–Hellman triples we are interested in tests for randomness other than simply considering uniform distribution.

(I) We may, for example, fix the k least (or most) significant bits and ask for uniform distribution of those. If $k < c \log m$ for a certain fixed constant $c(\delta)$ we are still able obtain this in the range $t > m^{3/4+\delta}$ provided that we assume $\Delta < m^{\delta}$, a condition which holds for most pairs (p, ℓ).

(II) Another natural problem to consider is multi-dimensional uniform distribution. Let

$$\Sigma = (\sigma_0, \dots, \sigma_{s-1})$$

denote s binary strings, each of length k, and let $L(\Sigma)$ be the number of $n \leq t$ such that, for each i, $i = 0, \dots, s-1$, σ_i is the string of the k least significant bits of u_{n+i}. We can study this via the multidimensional exponential sum

$$S_{\underline{a}} \doteq \sum_{n=1}^{t} e_m \left(\sum_{i=0}^{s-1} a_i u_{n+i} \right)$$

obtaining the following result.

THEOREM 6.5. *Assume $\Delta < m^{\delta}$ and $t > m^{3/4+\delta}$. Then, provided that*

$$s < c(\delta) \frac{\log m}{\log \log m},$$

we have equi-distribution of $L(\Sigma)$ amongst these s-tuples.

6.1. PERIOD OF THE BIT GENERATOR

Even the case $k = 1$ in the previous theorem has an interesting application, namely to the period of the "bit generator." This is the random number generator we obtain by taking the sequence of numbers produced, say by the power generator, and then throwing away all information other than the least significant bit. To fix our ideas, let us take the exponent $e = 2$, that is the Blum–Blum–Shub generator.

It is obvious that the bit generator so formed is also periodic and moreover that its period, say τ, is a divisor of the period t of the original Blum–Blum–Shub generator. In order for the bit generator to behave in a random unpredictable way we should like to know that the period τ is large. However, it could conceivably be the case that, even when t is large, its divisor τ is not. The previous theorem, however, rules out that possibility, at least to some extent.

Indeed, since there take place all possible patterns of s-tuples for consecutive occurrences of the last bit, with s as large as $s \approx c \log m / \log \log m$, it follows immediately that $\tau \geq 2^s > m^{c/\log\log m}$. Actually one can do a little bit better than that. Instead of counting the occurrence of these s-tuples with weight one we may use a standard trick and count them instead weighted by a smooth function (in fact the sharp-cornered Fejer kernel is smooth enough for this purpose). For the bit generator this gives, under the same assumptions $\Delta < m^{\delta}$ and $t > m^{3/4+\delta}$ as before, the bound $\tau > m^{\delta'}$, for some $\delta'(\delta) > 0$.

7. Large Periods and the Carmichael Function

Our work in this section is joint with Pomerance and Shparlinski (Friedlander et al., 2001a; Friedlander et al., 2001b). Some of the results have since been sharpened by Kurlberg and Pomerance (Kurlberg and Pomerance, 2005).

In the theorems from the previous sections we always required an assumption that, to some extent, the period t of our sequence was large. In any case it is intuitively clear, as was pointed out early in these notes, that we should want a large period.

For the discussion of large periods it is useful to recall the Carmichael function $\lambda(m)$, defined to be the least $e \in \mathbb{N}$ such that $a^e \equiv 1 \pmod m$ for all a with $(a, m) = 1$. This function is quite reminiscent of the more familiar Euler function $\varphi(m)$, specifically

$$\lambda(p^f) = \begin{cases} \frac{1}{2}\varphi(p^f) & p = 2,\ f \geq 3, \\ \varphi(p^f) & \text{else,} \end{cases}$$

but, rather than being multiplicative, it has the somewhat more awkward decomposition

$$\lambda(m) = \operatorname{lcm}\{\lambda(p_1^{f_1}), \dots, \lambda(p_\nu^{f_\nu})\}$$

for $m = p_1^{f_1} \cdots p_\nu^{f_\nu}$.

The Carmichael function has the following relation to the power generator, which, as we recall, was defined as

$$u_n \equiv u_{n-1}^e \pmod m, \qquad u_0 = \alpha.$$

If $(e, \lambda(m)) = 1$ then this is purely periodic and $t = \operatorname{ord}_s e$ where $s = \operatorname{ord}_m \alpha$. Thus, the largest possible value for t as α, e vary is $\lambda(\lambda(m))$. We would like to know that this is large.

An early result in this direction is due to Erdös, Pomerance and Schmutz (Erdös et al., 1991). They showed that $\lambda(n)$ is usually large, in particular that

$$\lambda(n) = n \exp(-\log_2 n \log_3 n - c \log_2 n + O(\log_2 n))$$

for all but $o(N)$ integers $n \leq N$. Here, and subsequently, $\log_k$ denotes the k-th iterate of the logarithm.

To get a handle on $\lambda(\lambda(p\ell))$ or even on $\lambda(\lambda(n))$ we need a result which holds apart from a smaller exceptional set and, as a result, have to settle for a somewhat weaker bound. We have

THEOREM 7.1. *Let N be large. Provided that $\Delta \geq (\log_2 N)^3$, the number of $n \leq N$ with*

$$\lambda(n) \leq n \exp(-\Delta)$$

is $\ll N \exp(-0.69(\Delta \log \Delta)^{1/3})$.

The proof of this is elementary but not particularly simple. The main ingredients are sieve bounds such as the Brun–Titchmarsh theorem and results on the distribution of smooth numbers.

This theorem is used to show that the iterated Carmichael function is also usually large. We obtain:

THEOREM 7.2. *Let Q be large and $\Delta \geq 2(\log_2 Q)^3$. The number $N(Q)$ of pairs p, ℓ of primes, $1 < p < \ell \leq Q$, having*

$$\lambda(\lambda(p\ell)) < Q^2 \exp(-\Delta)$$

satisfies the bound

$$N(Q) \ll Q^2 \exp(-0.16(\Delta \log \Delta)^{1/3}).$$

We give a very brief idea of the proof. First consider those pairs with greatest common divisor $(p-1, \ell-1) \geq D$. The number of these is bounded by $\sum_{d \geq D} \pi(Q; d, 1)^2$ which is small in terms of D by virtue of the Brun–Titchmarsh theorem.

On the other hand, for small $(p-1, \ell-1)$ we can give a bound on the multiplicity of λ, specifically, for given n, the multiplicity of $\lambda(p\ell) = n$ with $(p - 1, \ell - 1) \leq D$ is $\leq D\tau(n)$, where $\tau(n)$ as usual denotes the number of positive divisors of n. To see this recall that

$$\lambda(p\ell) = \operatorname{lcm}\{p-1, \ell-1\} = \frac{(p-1)(\ell-1)}{\gcd(p-1, \ell-1)}$$

and there are at most D choices for the gcd and at most $\tau(n)$ choices for $p-1$.

Actually, we have just given a bound for the multiplicity of λ and we really need a bound for the multiplicity of its iterate, but an elaboration of the above ideas carry over to $\lambda \circ \lambda$.

It is worth noting that, if we are willing to settle for fewer pairs p, ℓ, then we can get $\lambda(\lambda(p\ell))$ to be very large indeed, namely:

THEOREM 7.3. *We have $\lambda(\lambda(p\ell)) > c_1Q^2$ for more than $c_2Q^2/(\log Q)^4$ pairs of primes $1 < p < \ell < Q$.*

Actually, we can do still a little better by assuming a famous conjecture. We expect but do not know that there are infinitely many primes p such that $(p-1)/2 = q$ is prime and $(q-1)/2 = r$ is also prime. Taking two distinct such primes, say p and ℓ, we find that $\lambda(\lambda(p\ell))$ is essentially $p\ell/8$.

The proof of Theorem 7.3 is motivated by the above construction. We use a lower bound sieve and the Bombieri–Vinogradov theorem to construct many primes p such that $(p-1)/2$ has no small prime factors and then consider pairs p, ℓ of such primes. Using elementary arguments and some applications of the upper bound sieve we show that most of the products $p\ell$ we have constructed have sufficiently large Carmichael function. Although we do not match the constant $\frac{1}{8}$ that follows from the conjecture, we actually obtain a large number of integers of the required type using this unconditional argument.

So far, we know that the largest possible period is usually large. This is not quite what we want. We want the period itself to usually be large, so we next show that, for most choices, the period is not too far from the largest possible. For this we use:

LEMMA 7.4. *Let $M \in \mathbb{N}$, and $K > 1$ be real. Then,*

$$\#\left\{g, 1 \le g \le M, (g, M) = 1, \operatorname{ord}_M g \le \frac{\lambda(M)}{K}\right\} \le \varphi(M)\tau(\lambda(M))/K.$$

Combining this with a number of other lemmas of similar type, we obtain the following result.

THEOREM 7.5. *Let Q be large and $\Delta \ge 6(\log_2 Q)^3$. For all pairs of primes p, ℓ with $1 < p < \ell \le Q$ except at most $Q^2 \exp(-0.1(\Delta \log \Delta)^{1/3})$ of them, the following statement holds*:

For all pairs (α, e) with $1 \le \alpha < m$, $1 \le e \le \lambda(m)$, $(\alpha, m) = (e, \lambda(m)) = 1$ (where $m = p\ell$) except at most $m\lambda(m)\exp(-0.2\Delta)$ of them, the period t of the sequence (u_n) satisfies $t \ge Q^2\exp(-\Delta)$ (and hence by our earlier results (u_n) is uniformly distributed).

There are a lot of parameters floating around in the previous theorem. It would be nice to fix some of them; especially we would like to fix the exponent e. We can do this using sieve methods, but getting a weaker result.

THEOREM 7.6. *There are more than $cQ^2/(\log Q)^4$ pairs p, ℓ of primes $1 < p < \ell \le Q$ such that, for all $1 \le \alpha < m$, $(\alpha, m) = 1$, $m = p\ell$, apart from $m^{1-\delta(\varepsilon)}$ of them, the period t of (u_n) with $e = 2$ satisfies $t > Q^{1-\varepsilon}$.*

Here, the big advantage is that we have been able to fix the exponent e (the specific choice $e = 2$ was just for convenience). There are however, as compared to the earlier Theorem 7.5, a number of disadvantages. For one thing, we obtain 'lots' of integers $p\ell$ rather than 'most', and a larger exceptional set of m. These are not so bad but, more seriously. we get Q rather than Q^2 in the lower bound for t. This is too small for our application to the power generator (for which we would need an exponent bigger than $\frac{3}{2}$).

There are however other applications of the knowledge that we have a large period. We conclude this section by giving two of these.

(I) Cycling attack on the RSA cryptosystem.

Let's recall the description of the very famous RSA method. Bob wishes to be able to receive secret messages from anybody, just as if say he is Air Canada and would like anybody to be able to send their credit card number in complete safety. Bob publishes, for everyone to see, his modulus, an integer m. He has obtained his modulus by multiplying together two large primes $m = p\ell$, but these he keeps a secret. He also publishes his exponent e, an integer larger than 1, having first checked that e satisfies the coprimality condition $(e, \varphi(m)) = 1$. It is easy for him to check this condition using the Euclidean algorithm and the exclusive knowledge that $\varphi(m) = (p-1)(\ell-1)$. He also, quickly computes, using this knowledge and keeping the result secret, the multiplicative inverse $\bar{e}$ of e modulo $\varphi(m)$.

Now, along comes Alice, who has decided she wants to go to Hawaii and therefore needs to send Bob a message, an integer α (her credit card number). She enters it on the website which, before transmission, encrypts by raising α to the exponent e modulo m. To decode, Bob simply computes the least residue of $(\alpha^e)^{\bar{e}}$ modulo m, in other words, α.

The point of all this is that all of the above computations can be done quickly once the factorization of m is known, but that factorization is difficult for everybody but Bob who constructed m from it in the first place.

7.1. CYCLING ATTACK

Next, along comes Conrad who would like to get his hands on that credit card number. He notices that the power generator can be viewed as a sequence of RSA encryptions:

$$\alpha \to (\alpha^e \to \alpha^{e^2} \to \cdots \to \alpha^{e^{t-1}} \to \alpha^{e^t}),$$

say $u_0, \ldots u_t$. Conrad, being in possession of the data m, e and able to view the encrypted message α^e, is able to generate all of the above quantities in the parentheses, that is, everything following u_0. But, once he sees the repetition $u_t = u_1$ he can backtrack and deduce that $u_0 = \alpha$ is given by u_{t-1} .

Now, if t is small, this very primitive attack succeeds, but not otherwise. Thus, it follows from Theorem 7.5 that, for almost all moduli m and almost all inputs (α, e) this attack fails, and, from Theorem 7.6 that, even for a given exponent e, it fails for most messages α.

(II) Timed-release crypto.

Suppose Bob wants to broadcast a message, but only wants it to be read at time T. One thing he can do is is to encrypt the message μ with a key κ, wait, and then send κ at time T. Actually, he doesn't even need to encrypt, he can just wait and send μ at time T.

But what if T is years, or even centuries? Suppose for example he wants to leave behind the claim that he has proved the Riemann Hypothesis but doesn't want to be queried too closely about the details of the proof. Rivest, Shamir and Wagner have suggested the following solution to this problem (that of leaving the message, not of proving RH).

7.2. PROPOSED SOLUTION

Encrypt μ with a key κ. Take an RSA modulus $m = p\ell$, and integers θ, e, s, as usual, and then evaluate $U \equiv \theta^{e^s} \pmod{m}$, which, due to our knowledge of $\varphi(m)$, can be done quickly in two steps as follows:

(I) Find
$$f \equiv e^s \ (\operatorname{mod} \varphi(m)), \quad 0 \le f \le \varphi(m),$$
which can be done in $O(\log s)$ steps by repeated squaring.

(II) Next evaluate $\theta^{e^s} \equiv \theta^f \pmod{m}$, which can also be done quickly, in $O(\log m)$ steps.

Now, compute
$$L \equiv \kappa + \theta^{e^s} \pmod{m}, \quad 0 \le L < m.$$

We can now make public all of m, θ, e, s, L. To recover κ one needs θ^{e^s} and, as in the previous example, the question becomes: How large is the period? The theorems in this section imply that usually it is large enough to cause this to take a long time.

8. Exponential Sums to General Modulus

In this section we shall consider two different types of sum estimated in (Friedlander et al., 2002).

(A) Sums over $\mathcal{K}$-invariant sets.

In dealing with the exponential sums to general modulus which are relevant for many of our applications we can take advantage of a special structure occurring therein. Let $\mathcal{K} \subset \mathcal{U}_t$ (the units modulo t), and $\mathfrak{Z} = \{z_1, \ldots, z_t\} \subseteq \mathbb{Z}_t$. We shall say that $\mathfrak{Z}$ is $\mathcal{K}$-invariant if for every k in $\mathcal{K}$, the sequence $kz_1, \ldots, kz_t$ is a permutation of $z_1, \ldots, z_t$.

As before, let $\theta \in \mathcal{U}_m$ have order t. With the above setup we can obtain an upper bound for the sum

$$S_a(m, \mathfrak{Z}, t) = \sum_{j=0}^{T-1} e_m(a\theta^{z_j}).$$

For illustration we state two special but basic cases.

(I) We take, for a fixed positive integer n,

$$\mathcal{K} = \mathfrak{Z} = \{x^n \mid x \in \mathcal{U}_t\}.$$

In this case we obtain

THEOREM 8.1. $\sum_{x \in \mathcal{U}_t} e_m(a\theta^{x^n}) \ll_\varepsilon (a, m)^{1/8} t^{21/32} m^{5/16+\varepsilon}$.

(II) Let $g \in \mathcal{U}_t$ have multiplicative order T. Take

$$\mathcal{K} = \mathfrak{Z} = \{g^j \mid j = 0, 1, \ldots, T-1\}.$$

For this sum one has

THEOREM 8.2. $\sum_{j=0}^{T-1} e_m(a\theta^{g^j}) \ll_\varepsilon (a, m)^{1/8} T^{3/8} t^{9/32} m^{5/16+\varepsilon}$.

Thus, in particular, if $(a, m) \leq m^\varepsilon$ and $T \geq t^{1-\varepsilon}$, the above result is non-trivial for $t \geq m^{10/11+\varepsilon}$. We next consider an application of this bound.

8.1. DISCREPANCY OF THE POWER GENERATOR

Here, we have an integer modulus $m \geq 2$, a seed θ with $(\theta, m) = 1$, and an exponent $g \geq 2$. We recall that the power generator is the sequence given by:

$$u_0 = \theta, \quad u_j \equiv u_{j-1}^g \pmod{m}$$
$$0 \leq u_j \leq m-1, \quad j = 1, 2, \ldots.$$

If, in particular, $\theta \in \mathcal{U}_m$ has multiplicative order t and $g \in \mathcal{U}_t$ has multiplicative order T then $u_j \equiv \theta^{g^j} \pmod{m}$, $j = 0, 1, \ldots$ is purely periodic with period T. From Theorem 8.2 and Lemma 2.3 we can bound the discrepancy of the power generator to an arbitrary modulus.

THEOREM 8.3. *Let $\mathcal{D}_m(t,T)$ denote the discrepancy of the sequence of fractional parts*

$$\left\{\frac{u_j}{m}\right\}, \quad j = 0, 1, \dots, T-1.$$

Then we have

$$\mathcal{D}_m(t,T) \underset{\varepsilon}{\ll} T^{-5/8} t^{9/32} m^{5/16+\varepsilon}.$$

If, in particular $t > m^{1-\varepsilon}$ then this is non-trivial for $T > m^{19/20+\varepsilon}$. If in addition, $T > m^{1-\varepsilon}$ then $\mathcal{D}_m \ll T^{-1/32+\varepsilon}$. In the special case of an 'RSA' modulus $m = p\ell$ one can get the stronger bound $T^{-1/8+\varepsilon}$.

(B) Double sums over general sets

We now consider, again for general modulus m, the sum

$$S_a(m,t,\mathcal{X},\mathcal{Y}) = \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} e_m(a\theta^{xy})$$

for arbitrary $a \in \mathbb{Z}_m$, $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{Z}_t$.

Here, no special structure is required. Because we have a double sum we can still get results if the sets are not too thin by means of a judicious use of Cauchy's inequality. Specifically, we may, for example, write

$$|S_a(m,t,\mathcal{X},\mathcal{Y})|^2 \le |\mathcal{X}| \sum_{x\in\mathcal{X}} \left| \sum_{y\in\mathcal{Y}} e_m(a\theta^{xy}) \right|^2 .$$

In the latter sum we may by positivity increase the summation over x to all of $\mathbb{Z}_t$, then expand the square and interchange the order of summation bringing the sum over x inside. This inner sum is now quite simple and we can proceed. One such result obtained is the following.

THEOREM 8.4. *We have*

$$S_a(m,t,\mathcal{X},\mathcal{Y}) \underset{\varepsilon}{\ll} |\mathcal{X}|^{1/2} |\mathcal{Y}|^{21/32} (a,m)^{1/8} t^{1/2} m^{5/16+\varepsilon}.$$

We note that $1/2 + 21/32 + 0 + 1/2 + 5/16 = 2 - 1/32$ and since this is less than two we obtain a nontrivial bound exhibiting cancellation in the sum provided the sets are sufficiently dense and the period is sufficiently large. We next consider an application of this result, given in (Friedlander et al., 2002).

8.2. COMMUNICATION COMPLEXITY OF THE DIFFIE–HELLMAN BIT

We consider the set of n-bit integers

$$\mathcal{B} = \{x \in \mathbb{Z} \mid 0 \le x \le 2^n - 1\}.$$

There is an obvious bijection between n-bit integers and their binary expansion, given by

$$x \in \mathcal{B} \leftrightarrow \underline{x} = (x_1, \dots, x_n) \in \mathbb{Z}_2^n.$$

Let $\underline{y} = (y_1, \dots, y_n)$ co-ordinatize a second such integer.

Now, let $f(\underline{x}, \underline{y}) = f(x_1, \dots, x_n, y_1, \dots, y_n)$ be a given function of $2n$ variables. Assume that we have two collaborating parties, one knowing $\underline{x}$ and the other knowing $\underline{y}$. Our goal is to create a "communication protocol" P such that for any inputs $x, y \in \mathcal{B}$, at the end one party is able to compute $f(\underline{x}, \underline{y})$.

For a given protocol P (that is an algorithm for exchanging the information), we define ψ_P: to be the largest number of bit exchanges required to compute $f(\underline{x}, \underline{y})$, taken over all inputs $x, y \in \mathcal{B}$. Then we define $\psi(f)$, the communication complexity of the function f, to be the minimum of ψ_P, taken over all possible protocols P.

A trivial upper bound for this complexity is given by $\psi(f) \le n$ since one party can simply tell everything to the other. It is quite a common phenomenon in complexity theory that lower bounds are much harder to obtain than upper bounds. Thus, in this case, to obtain an upper bound one only needs to examine the results from a single protocol. On the other hand, to give a lower bound one needs to say something about every conceivable protocol.

Given $x, y \in \mathcal{B}$ we study the communication complexity of the Diffie–Hellman key, in particular, of the Diffie–Hellman bit operation. By this is meant we take the sequence given by the common key and throw away all information except the last bit.

Specifically, for an odd integer m, and θ of multiplicative order t, choose n to be the largest integer with $2^n \le t$ and, for $x, y \in \mathcal{B}$, define

$$f(x_1, \dots, x_n, y_1, \dots, y_n) = \begin{cases} 1 & \text{if } \theta^{xy} \in \{1, 3, 5, \dots, m-2\}, \\ 0 & \text{if } \theta^{xy} \in \{2, 4, 6, \dots, m-1\}. \end{cases}$$

As a consequence of the previous Theorem 8.4 one may deduce a reasonable lower bound for the complexity of this function.

THEOREM 8.5. *Let m be an odd integer, $\delta > 0$ be real, and assume the period t of θ modulo m satisfies $t \ge m^{10/11+\delta}$. Then, the communication complexity of the Diffie–Hellman bit operation satisfies the bound*

$$\psi(f) \ge \Big(\frac{11}{32}\delta - o(1)\Big)n.$$

This lower bound provides, a fortiori, a lower bound for the communication complexity of the Diffie–Hellman secret key θ^{xy} itself. Moreover, in view of the trivial bound $\psi(f) \leq n$, this bound is rather good; only the constant is in doubt. The above theorem depends heavily on an interesting relationship between the communication complexity and the "combinatorial" discrepancy established in (Babai et al., 1992).

We define the combinatorial discrepancy $\Delta(f)$ of f by

$$\Delta(f) = 2^{-2n} \max_{\mathcal{X},\mathcal{Y}\subseteq\mathcal{B}} |N_0(\mathcal{X}, \mathcal{Y}) - N_1(\mathcal{X}, \mathcal{Y})|,$$

where $N_r(\mathcal{X}, \mathcal{Y})$ denotes, for $r = 0, 1$, the number of pairs $x \in \mathcal{X}$, $y \in \mathcal{Y}$ with $f(\underline{x}, y) = r$.

Theorem 8.4 can be applied in a way now familiar to give a bound for the combinatorial discrepancy. On the other hand, as a part of Lemma 2.2 of (Babai et al., 1992) we have the elegant inequality

$$\Delta(f)2^{\psi(f)} \geq 1,$$

from which Theorem 8.5 follows.

9. Sums over Elliptic Curves

Many of the most useful and important cryptographic constructs take place within the group $\mathbb{F}_p^\times$ but make some sense within virtually any finite Abelian group. On the other hand, for some groups, say for example the additive group $\mathbb{F}_p$, the description of the group is too simple and useful applications are lacking.

For the past twenty years, since the work of N. Koblitz and of V. Miller, there has been an increasing interest in the cryptographic uses of the finite group of points of an elliptic curve over a finite field. In this section we give results from (Banks et al., 2006a) which deal with analogues to some of the topics discussed earlier, but now within this elliptic setting.

Let $\mathcal{E}$ be an (ordinary) elliptic curve with Weierstrass equation $Y^2 = X^3 + AX + B$ over the finite field $\mathbb{F}_p$. Denote a typical point on the curve, and its coordinates by $Q = (x(Q), y(Q))$. Let G be a fixed point of order t on the curve $\mathcal{E}$. Analogous to the sum $S_a(m, t, \mathcal{X}, \mathcal{Y})$ considered in the previous section, we study the sums

$$W_a(\mathcal{U}, \mathcal{V}) = \sum_{u\in\mathcal{U}} \sum_{v\in\mathcal{V}} \alpha_u \beta_v e_p(ax(uvG))$$

where $a \in \mathbb{F}_p^\times$ is fixed, $\mathcal{U}, \mathcal{V} \subseteq \mathbb{Z}_t$ are given subsets, and α_u, β_v are complex coefficients. For this sum we obtain the following result.

THEOREM 9.1. *Assume* $|\alpha_u| \le 1$, $|\beta_v| \le 1$ *and* $\varepsilon > 0$. *Then*

$$W_a(\mathcal{U}, \mathcal{V}) \underset{\varepsilon}{\ll} t^{5/6}(\#\mathcal{U}\#\mathcal{V})^{1/2} p^{1/12+\varepsilon}.$$

We note in particular that, if $\mathcal{U}, \mathcal{V}$ are almost dense in $\mathbb{Z}_t$ in the sense that $\#\mathcal{U}, \#\mathcal{V} > t^{1-o(1)}$ then the theorem provides a non-trivial bound in the range $t > p^{1/2+\varepsilon}$. This range, which is wider than those given earlier, can be attributed to the incorporation of ideas of Garaev from (Garaev, 2005). In this elliptic case the proof eventually reduces to estimates for exponential sums over finite fields in which an earlier reliance on Weil's bound is replaced by bounds due to Bombieri (Bombieri, 1966). We remark that one obtains the same result if, in the sum $W_a(\mathcal{U}, \mathcal{V})$, we replace the x-coordinate in the exponential by the y-coordinate or indeed by any nonconstant rational function in x and y.

In the same fashion as earlier, the above theorem implies bounds for the exponential sum

$$\sum_{u=1}^{t} \sum_{v=1}^{t} e_p(ax(uG) + bx(vG) + cx(uvG)),$$

and hence uniformity of distribution results for Diffie–Hellman triples on an elliptic curve, that is the triples

$$(x(uG), x(vG), x(uvG)).$$

As a less immediate application, one can consider the analogue of the power generator on $\mathcal{E}$. Again, let G be a fixed point on $\mathcal{E}$ having order t, fix an integer $e \ge 2$ such that $(e, t) = 1$. Define the elliptic power generator to be the sequence:

$$U_n = eU_{n-1}, \quad n = 1, 2, \ldots,$$

beginning with the initial condition $U_0 = G$. The previous theorem implies, by extending the method of (Friedlander et al., 1999) which gave Theorem 6.1, the following bound for the pertinent exponential sum.

THEOREM 9.2. *We have, for* $a \in \mathbb{F}_p^\times$,

$$\sum_{n \le N} e_p(ax(U_n)) \underset{\varepsilon}{\ll} t^{5/9} N^{1/3} p^{1/18+\varepsilon}.$$

This in turn gives the following uniformity of distribution result.

COROLLARY 9.3. *Let T be order of e modulo t. If*

$$T \geq N \geq p^{11/12+\varepsilon},$$

then the sequence of fractional parts

$$\left\{\frac{x(U_n)}{p}\right\}, \quad n = 0, 1, \ldots, N,$$

is uniformly distributed.

10. Proof Sketch of Theorem 4.1

The ideas in this section originate in (Canetti et al., 1999) and (Canetti et al., 2000). Recall that our goal is an upper bound for the sum

$$V_{a,c}(t) = \sum_{y=1}^{t} \left| \sum_{x=1}^{t} e_p(a\theta^x + c\theta^{xy}) \right|^4,$$

where t is the multiplicative order of $\theta \pmod p$. Using an idea exploited in some of our earlier arguments, we note that the sum is invariant under the translation of variables $x \to x + z$. This leads us to the following estimation:

$$\begin{aligned} V_{a,c}(t) &= t^{-1} \sum_{y=1}^{t} \sum_{z=1}^{t} \left| \sum_{x=1}^{t} e_p(a\theta^{x+z} + c\theta^{(x+z)y}) \right|^4 \\ &= t^{-1} \sum_{y=1}^{t} \sum_{z=1}^{t} \left| \sum_{x=1}^{t} e_p(a\theta^{z}\theta^{x} + c\theta^{zy}\theta^{xy}) \right|^4 \\ &\leq t^{-1} \sum_{y=1}^{t} \sum_{\lambda,\mu=0}^{p-1} \left| \sum_{x=1}^{t} e_p(\lambda\theta^{x} + \mu\theta^{xy}) \right|^4, \end{aligned}$$

where, in the last step we have used the fact that, for each y, the pairs $(a\theta^z, c\theta^{zy})$ are distinct modulo p as z runs through the values $z = 1, 2, \ldots, t$. Thus,

$$V_{a,c}(t) \leq t^{-1}p^2T,$$

where T is the number of solutions of the system of congruences

$$\begin{cases} \theta^{x_1} + \theta^{x_2} \equiv \theta^{x_3} + \theta^{x_4} & \pmod p, \\ \theta^{x_1y} + \theta^{x_2y} \equiv \theta^{x_3y} + \theta^{x_4y} & \pmod p, \end{cases} \tag{1}$$

as the variables range over $x_1, x_2, x_3, x_4, y = 1, 2, \ldots, t$.

Write $\theta = g^m$ where g is a primitive root modulo p so that m is given by $p-1 = mt$. For a given quadruple (x_1, x_2, x_3, x_4) satisfying the first congruence, the number of y satisfying the second is just N/m where N is the number of solutions z of

$$z^{k_1} + z^{k_2} - z^{k_3} - z^{k_4} \equiv 0 \pmod{p}$$

as $z = 1, 2, \dots, p-1$, and where $k_i = mx_i$. We need a bound for the number of such z and this leads us to the following problem.

10.1. NUMBER OF ROOTS OF A SPARSE POLYNOMIAL OVER A FINITE FIELD

Although we are only going to use the result for fields of prime order, it is no extra work to be somewhat more general. Let, as usual, $\mathbb{F}_q$ denote the finite field of q elements.

LEMMA 10.1. *Let $a_1, a_2, \dots, a_n$ be elements of $\mathbb{F}_q^\times$ and $k_1, \dots, k_n \in \mathbb{Z}$. Then, the number N of solutions of the equation*

$$\sum_{i=1}^{n} a_i z^{k_i} = 0,$$

as z ranges through $\mathbb{F}_q^\times$, satisfies the bound

$$N \le 2^{1-1/(n-1)} D^{1/(n-1)} + O(q^{1-2/(n-1)} D^{2/(n-1)}),$$

where

$$D = \min_{1 \le i \le n} \max_{j \ne i}(k_j - k_i, q-1).$$

REMARK. This bound is suited toward the situation where we are dealing with a *sparse* polynomial; that is it is good in case the number n of monomials is small. The degree of the polynomial is not that important except in the case where D is large. Some restriction of this type is certainly necessary since, for example, for odd q the polynomial $z^{(q-1)/2} - 1$ is very sparse but has many roots, $(q-1)/2$ of them.

The general idea behind the proof of the lemma is as follows. We use the box principle to choose a small positive integer ℓ, which is less than some quantity L, the best choice of which will turn out to depend on D, but is such that, when we make a change of variable replacing z by z^ℓ, the degrees k_i of the monomials are replaced by $k_i\ell \equiv m_i \pmod{q-1}$ where the new degrees m_i are not too large, say $\le M$. Let us assume that $(\ell, q-1) = 1$ so that no solutions are lost in the change of variables; if there is a common divisor the situation is a little more complicated.

By Fermat's little theorem, replacing the exponents $k_i\ell$ by m_i, we still obtain the same solutions. Since our polynomial is sparse, there are only a few congruence conditions we need to satisfy, that is only a few m_i that need to be small. Finally, since the resulting polynomial has small degree, no more than M, there is an acceptable bound for the number of roots.

10.2. ANCHORS AND PAIRINGS

We are now ready to return to the proof of Theorem 4.1. Recall that we were considering the system

$$\begin{aligned} \theta^{x_1} + \theta^{x_2} &\equiv \theta^{x_3} + \theta^{x_4} \pmod p, \\ \theta^{x_1 y} + \theta^{x_2 y} &\equiv \theta^{x_3 y} + \theta^{x_4 y} \pmod p. \end{aligned}$$

We need to input our bound for the number of y coming from Lemma 10.1 and then sum over the quadruples (x_1, x_2, x_3, x_4) satisfying the first congruence.

Note that the bound in the lemma depends on the quadruple only to the extent of the value of D. Hence we require, for each positive integer d, an upper bound for the number of quadruples (x_1, x_2, x_3, x_4) satisfying

$$\theta^{x_1} + \theta^{x_2} \equiv \theta^{x_3} + \theta^{x_4} \pmod p, \qquad 1 \le x_i \le t,$$

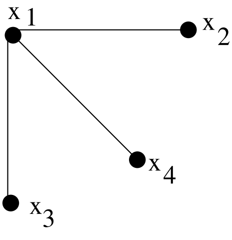
with

$$\min_{1\le i\le 4} \max_{j\ne i} (x_i - x_j, t) = d.$$

Actually, we lose virtually nothing by replacing the equality in the last condition by an inequality, which makes the counting easier. Suppose that, for each i, $1 \le i \le 4$, there exists a $j \ne i$ such that $(x_i - x_j, t) \ge d$.

There are two types of minimal configuration, unique apart from permutation of the indices, which can bring this about. It may be that one of the variables, say x_1, anchors the others as described by the following diagram.

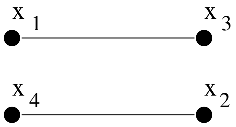
(I) Anchor



Here, we have joined the vertices x_i, x_j by an edge in case $(x_i - x_j, t) \geq d$, and if there are other edges we just ignore them.

In the second case, the variables may pair off and, using the same description, we have

(II) Pairing



We need upper bounds for the number of occurrences of each.

10.3. BOUNDS FOR THE NUMBER OF PAIRINGS

It seems intuitively obvious that the number of anchors will be smaller than the number of pairings and can be ignored in this sketch. Indeed, it does turn out that the same ideas we use for bounding the number of pairings can be used to give (even better) bounds for the anchors.

Fix coefficients $a_1, a_2, a_3, a_4 \not\equiv 0 \pmod p$ and divisors $d_1 \mid t$, $d_2 \mid t$. Let $Q_{d_1,d_2}(t)$ denote the number of solutions of the congruence

$$a_1\theta^{x_1} + a_2\theta^{x_2} + a_3\theta^{x_3} + a_4\theta^{x_4} \equiv 0 \pmod p,$$

as the variables run through the range

$$1 \leq x_1, x_2, x_3, x_4 \leq t,$$

subject to the constraints

$$x_1 \equiv x_3 \pmod{d_1}, \quad x_2 \equiv x_4 \pmod{d_2}.$$

We give two different bounds for this number.

(A) Elementary counting bound.

LEMMA 10.2.

$$Q_{d_1,d_2}(t) \leq \frac{t^3}{d_1 d_2} + t^2.$$

Proof. First, we count those quadruples for which we have the additional condition

$$a_1\theta^{x_1} + a_3\theta^{x_3} \equiv 0 \pmod p.$$

Each x_1 determines (at most) one value of x_3. Then, each x_2 determines (at most) one value of x_4. Hence, there at most t^2 of these quadruples.

Next, we count the others. There are no more than t^2/d_1 choices for the pair x_1, x_3 and once that pair is determined it fixes a single non-zero value for the class of

$$\theta^{x_2}(a_2 + a_4\theta^{x_4 - x_2}).$$

The choice of $x_4 - x_2$, which can be made in precisely t/d_2 different ways, then determines the rest. □

(B) Exponential sum bound. Replacing the above simple counting with an argument using exponential sums, we obtain a result which is better than (A) if d_1, d_2 are small.

LEMMA 10.3.

$$Q_{d_1,d_2}(t) = \frac{t^4}{d_1 d_2 p} + O(tp).$$

To prove this, we detect the condition

$$x_1 \equiv x_3 \pmod{d_1}$$

using additive characters, that is

$$d_1^{-1} \sum_{j=0}^{d_1 - 1} e\left(\frac{j(x_1 - x_3)}{d_1}\right),$$

and similarly, each of the conditions $x_2 \equiv x_4 \pmod{d_2}$ and

$$a_1\theta^{x_1} + \cdots + a_4\theta^{x_4} \equiv 0 \pmod{p}.$$

Inserting these, we are led to an exponential sum in seven variables, but not much is required for its estimation. It turns out that the ingredients we need are bounds for the sum

$$\sigma(a, b) = \sum_{x=1}^{t} e_p(a\theta^x) e_{p-1}(bx),$$

which was already encountered in Section 3 in connection with the discrete logarithm problem.

In particular, to complete the proof of the lemma we require the following simple facts about this sum.

LEMMA 10.4.

(1) *If* $p \nmid a$ *we have*

$$|\sigma(a, b)| \leq p^{1/2}.$$

(2) *We have the mean square evaluation*

$$\sum_{a=1}^{p} |\sigma(a,b)|^2 = tp.$$

The first of these is a quite old Gauss sum bound, known certainly to Korobov. The second is also quite old and follows at once on interchanging the order of summation and using the basic orthogonality property of additive characters.

We have now assembled all of the pieces. Making the optimal choice between Lemmas 10.2 and 10.3, we estimate the outer sum over the quadruples (x_1, x_2, x_3, x_4) giving in turn a bound for the number T of solutions to the system (1) of congruences and thereby to the sum $V_{a,c}(t)$, leading to Theorem 4.1.

Acknowledgements

The author's work on these topics is, as indicated in the bibliography, joint with a number of co-authors. All of these, Igor Shparlinski especially, deserve mention. During part of this time the work received the support of a Killam Research Fellowship, and throughout all of this time (indeed, twenty-five years as of April 2006) received partial support from NSERC grant A5123. These notes were to a large extent written during an enjoyable stay, with partial support, at the Centre de Recherches Mathématiques in Montréal, during their special semester on *Analytic Number Theory*, and benefited from the helpful criticisms of Andrew Granville.

References

Babai, L., Nisan, N., and Szegedy, M. (1992) Multiparty protocols, pseudorandom generators for logspace and time-space trade-offs, *J. Comput. System Sci.* **45**, 204–232.

Banks, W. D., Friedlander, J. B., Garaev, M., and Shparlinski, I. E. (2006a) Double character sums over elliptic curves and finite fields, *Pure Appl. Math. Q.*, J. Coates 60th birthday volume.

Banks, W. D., Friedlander, J. B., Konyagin, S. V., and Shparlinski, I. E. (2006b) Incomplete exponential sums and Diffie–Hellman triples, *Math. Proc. Cambridge Philos. Soc* **40**, 193–206.

Bombieri, E. (1966) On exponential sums in finite fields, *Amer. J. Math.* **88**, 71–105.

Bourgain, J. (2004) Estimates on exponential sums related to Diffie–Hellman distributions, *Comptes Rendus Mathématique* **338**, 825–830.

Bourgain, J. (2005) Estimates on exponential sums related to the Diffie–Hellman distributions, *Geom. Funct. Anal.* **15**, 1–34.

Canetti, R., Friedlander, J. B., Konyagin, S., Larsen, M., Lieman, D., and Shparlinski, I. E. (2000) On the statistical properties of Diffie–Hellman distributions, *Israel J. Math.* **120**, 23–46.

Canetti, R., Friedlander, J. B., and Shparlinski, I. E. (1999) On certain exponential sums and the distribution of Diffie–Hellman triples, *J. London Math. Soc.* **59**, 799–812.

Erdös, P., Pomerance, C., and Schmutz, E. (1991) Carmichael's lambda function, *Acta Arith.* **58**, 363–385.

Friedlander, J. B., Konyagin, S. V., and Shparlinski, I. E. (2002) Some doubly exponential sums over Z_m, *Acta Arith.* **105**, 349–370.

Friedlander, J. B., Lieman, D., and Shparlinski, I. E. (1999) On the distribution of the RSA generator, In *Sequences and their applications*, Springer, London, pp. 205–212, Discrete Math. Theor. Comput. Sci.

Friedlander, J. B., Pomerance, C., and Shparlinski, I. E. (2001a) Period of the power generator and small values of Carmichael's function, *Math. Comp.* **70**, 1591–1605, Corrigendum, *ibid.* **71**, 1803–1806.

Friedlander, J. B., Pomerance, C., and Shparlinski, I. E. (2001b) Small values of the Carmichael function and cryptographic applications, In *Proc. Conf. Computational Number Theory and Cryptography*, Vol. 20 of *Progress in Computer Science and Logic*, pp. 25–32, Birkhäuser.

Friedlander, J. B. and Shparlinski, I. E. (2001) On the distribution of the power generator, *Math. Comp.* **70**, 1575–1589.

Garaev, M. Z. (2005) Double exponential sums related to Diffie–Hellman distributions, *Int. Math. Res. Not.* **17**, 1005–1014.

Kurlberg, P. and Pomerance, C. (2005) On the periods of the linear congruential and power generators, *Acta Arith* **119**, 149–169.

Shparlinski, I. E. (2002) On the distribution of the Diffie–Hellman pairs, *Finite Fields Appl.* **8**, 131–141.

THE DISTRIBUTION OF PRIME NUMBERS

K. Soundararajan
University of Michigan

What follows is an expanded version of my lectures at the NATO School on Equidistribution. I have tried to keep the informal style of the lectures. In particular, I have sometimes oversimplified matters in order to convey the spirit of an argument.

1. The Cramér Model and Gaps Between Consecutive Primes

The prime number theorem tells us that $\pi(x)$, the number of primes below x, is $\sim x/\log x$. Equivalently, if p_n denotes the n-th smallest prime number then $p_n \sim n \log n$. What is the distribution of the gaps between consecutive primes, $p_{n+1} - p_n$?

We have just seen that $p_{n+1} - p_n$ is approximately $\log n$ "on average." How often do we get a gap of size $2 \log n$, say; or of size $\frac{1}{2} \log n$? One way to make this question precise is to fix an interval $[\alpha, \beta]$ (with $0 \le \alpha < \beta$) and ask for

$$\lim_{N\to\infty} \frac{1}{N} \# \Big\{ 2 \le n \le N : \frac{p_{n+1} - p_n}{\log n} \in [\alpha, \beta] \Big\}. \tag{1}$$

Does this limit exist, and if so what does it equal?

Here is another way to formulate this question. Consider intervals of the form $[n, n + \log n]$ as n ranges over integers up to N. On average such an interval contains one prime. But of course some intervals may not contain any prime, and others may contain several. Given a non-negative integer k, how often does such an interval contain exactly k primes? What is

$$\lim_{N\to\infty} \frac{1}{N} \#\{ n \le N : \pi(n + \log n) - \pi(n) = k \}? \tag{2}$$

Or more generally, for a fixed real number $\lambda > 0$ we may ask for

$$\lim_{N\to\infty} \frac{1}{N} \#\{ n \le N : \pi(n + \lambda \log n) - \pi(n) = k \}? \tag{3}$$

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 59–83.

In this lecture we will describe the conjectured answers to these questions, but we confess at the outset that no one knows how to prove those conjectures. While conjecturing the prime number theorem, Gauss stated that the 'density of primes' around x should be $1/\log x$. He based his conjecture on extensive numerical investigations. In particular he divides the numbers up to three million into intervals of length 100 (a "centad") and meticulously tabulates the number of centads with no primes, exactly one prime etc.[1] While he does not seem to make a synthesis of his results (except to conjecture the prime number theorem) it seems clear that he was seeking to understand a question like (3). It was left to Harald Cramér (Cramér, 1936) to set Gauss's work on a probabilistic footing.

CRAMÉR'S MODEL. *The primes behave like independent random variables $X(n)$ $(n \geq 3)$ with $X(n) = 1$ (the number n is 'prime') with probability $1/\log n$, and $X(n) = 0$ (the number n is 'composite') with probability $1 - 1/\log n$.*

Let us suppose that the primes behave like a typical sequence in this random model, and answer questions (1) and (3). We want the 'probability' that $p_{n+1} - p_n$ lies between $\alpha \log n$ and $\beta \log n$. Thus, given the prime p_n, we want $p_n + 1, \ldots, p_n + h - 1$ to be composite, and $p_n + h$ to be prime, where $\alpha \log n \leq h \leq \beta \log n$. According to Cramér's model, this occurs with probability

$$\sum_{\alpha \log n \leq h \leq \beta \log n} \prod_{j=1}^{h-1} \left(1 - \frac{1}{\log(p_n + j)}\right) \frac{1}{\log(p_n + h)} \sim \sum_{\alpha \log n \leq h \leq \beta \log n} \left(1 - \frac{1}{\log n}\right)^{h-1} \frac{1}{\log n}$$

since $\log(p_n + j) \sim \log n$ as $p_n \sim n \log n$ and $j \ll \log n$. This is

$$\sim \sum_{\alpha \leq h/\log n \leq \beta} e^{-h/\log n} \frac{1}{\log n} \sim \int_{\alpha}^{\beta} e^{-t}\, dt,$$

for large n, because the LHS looks like a Riemann sum approximation to the integral in the RHS. This is the conjectured answer to question (1): the probability "density" of finding $p_{n+1} - p_n$ close to $t \log n$ is e^{-t}. This is an example of what is known as a "Poisson process" in the probability literature, see Feller (Feller, 1966).

EXERCISE 1.1. Show similarly that the Cramér model predicts that the answer to question (3) is $\lambda^k e^{-\lambda}/k!$. This is the Poisson distribution with parameter λ.

[1] We refer the reader to www.math.princeton.edu/~ytschink/.gauss for scans of Gauss's manuscripts showing these calculations.

The reader may well object that these predictions are dubious: clearly the probability that n and $n+1$ are both primes must be zero, but the Cramér model assigns this event a probability $1/(\log n \log(n+1))$. More generally, suppose we are given a set $\mathcal{H} = \{h_1, \ldots, h_k\}$ of k distinct integers, and we ask for the number of integers $n \leq x$ with $n+h_1$, $n+h_2, \ldots, n+h_k$ all being prime. The Cramér model would predict an answer of $\sim x/(\log x)^k$, but clearly we must take into account arithmetic properties of the set $\mathcal{H}$. For example, if there were a prime p such that the integers $h_1, \ldots, h_k$ occupied all the residue classes (mod p) then the integers $n+h_1, \ldots, n+h_k$ would also occupy all the residue classes (mod p). In particular one of these numbers would be a multiple of p, and so there can only be finitely many values of n with $n+h_1, \ldots, n+h_k$ all being prime.

In (Hardy and Littlewood, 1922) Hardy and Littlewood proposed the prime k-tuple conjecture that

$$\#\{n \leq x : n+h_1, n+h_2, \ldots, n+h_k \text{ prime}\} \sim \mathfrak{S}(\mathcal{H}) \frac{x}{(\log x)^k}, \tag{4}$$

for a certain constant $\mathfrak{S}(\mathcal{H})$ called the 'singular series.' The constant $\mathfrak{S}(\mathcal{H})$ equals 0 if the elements $h_1, \ldots, h_k$ occupy a complete set of residue classes (mod p) for some prime p, and $\mathfrak{S}(\mathcal{H})$ is positive otherwise. We will describe this conjecture in more detail below. The aim of this lecture is to describe a beautiful calculation of Gallagher (Gallagher, 1976) which shows that the Hardy–Littlewood conjecture (4) implies the same distribution of gaps between primes predicted by the Cramér random model. The crux of his proof is that although $\mathfrak{S}(\mathcal{H})$ is not always 1 (as the Cramér model would have), it is ~ 1 on average over all k-element sets $\mathcal{H}$ with the $h_j \leq h$.

1.1. THE HARDY–LITTLEWOOD CONJECTURE

We now motivate the Hardy–Littlewood conjecture (4) and describe the singular series $\mathfrak{S}(\mathcal{H})$ that arises there. As a toy model for prime numbers let us fix an integer q and consider the reduced residue classes (mod q). Out of the q total residue classes, there are $\phi(q)$ reduced classes, and we may think of $\phi(q)/q$ as the 'probability' of a class being reduced. Now suppose we are given the set $\mathcal{H} = \{h_1, \ldots, h_k\}$ and we ask for the number of $n \pmod q$ such that $n+h_1, \ldots, n+h_k$ are all coprime to q. For convenience, let us just think of square-free q. If these k events were independent then the answer would be $q(\phi(q)/q)^k$. The correct answer is a little different: for each prime p that divides q we need n to avoid the residue classes $-h_1, \ldots, -h_k$ (mod p). Let $\nu_{\mathcal{H}}(p)$ denote the number of distinct residue classes occupied by $\mathcal{H} \pmod p$. Thus n must lie in one of $p - \nu_{\mathcal{H}}(p)$ residue classes (mod p).

Using the Chinese remainder theorem we see easily that the correct answer is

$$\prod_{p|q}(p-\nu_{\mathcal{H}}(p)) = q\prod_{p|q}\Big(1-\frac{\nu_{\mathcal{H}}(p)}{p}\Big) = q\Big(\frac{\phi(q)}{q}\Big)^k\prod_{p|q}\Big(1-\frac{\nu_{\mathcal{H}}(p)}{p}\Big)\Big(1-\frac{1}{p}\Big)^{-k}.$$

Let us write $\mathfrak{S}(\mathcal{H};q) = \prod_{p|q}(1-\nu_{\mathcal{H}}(p)/p)(1-1/p)^{-k}$. We have seen that the answer for the number of $n \pmod q$ with $n+h_1,\dots,n+h_k$ all being coprime to q involves correcting the guess $q(\phi(q)/q)^k$ by the factor $\mathfrak{S}(\mathcal{H};q)$ which keeps track of the arithmetic properties of the set $\mathcal{H}$. Now let us consider what happens when we take $q = q_\ell = \prod_{p\le\ell} p$ and let ℓ go to infinity. As $\ell \to \infty$ we see that

$$\mathfrak{S}(\mathcal{H};q_\ell) \to \prod_p\Big(1-\frac{\nu_{\mathcal{H}}(p)}{p}\Big)\Big(1-\frac{1}{p}\Big)^{-k}.$$

The infinite product above converges because if p is larger than all the h_j's then $\nu_{\mathcal{H}}(p) = k$ and so $(1-\nu_{\mathcal{H}}(p)/p)(1-1/p)^{-k} = (1-k/p)(1-1/p)^{-k} = 1+O(p^{-2})$. This infinite product is the singular series[2]:

$$\mathfrak{S}(\mathcal{H}) := \prod_p\Big(1-\frac{\nu_{\mathcal{H}}(p)}{p}\Big)\Big(1-\frac{1}{p}\Big)^{-k}. \tag{5}$$

Further if $n+h_1,\dots,n+h_k$ are coprime to q_ℓ with ℓ large, then they have no small prime divisors, and may reasonably be viewed as a kind of approximation to primes. Thus in formulating a conjecture on the number of $n \le x$ with $n+h_1,\dots,n+h_k$ being prime, a natural guess is to take the random answer $x/(\log x)^k$, and multiply it by the arithmetical correction factor $\mathfrak{S}(\mathcal{H})$. This is precisely the Hardy–Littlewood conjecture (4). It is immediate from (5) that $\mathfrak{S}(\mathcal{H}) = 0$ if and only if $\mathcal{H}$ exhausts a compete set of residue classes $\pmod p$ for some p.

1.2. GALLAGHER'S CALCULATION

We will now describe Gallagher's argument, using the Hardy–Littlewood conjecture (4) to justify the distribution of gaps between primes predicted by the Cramér model. The precise problem we consider is a close variant of question (3). Let λ be a positive real number, and let N be large. We set $h = \lambda\log N$ and seek to understand the distribution of $\pi(n+h)-\pi(n)$ as n varies over the natural numbers below N. To understand this quantity, we

[2] The term arises from Hardy and Littlewood's original derivation of their conjecture using the circle method. Here $\mathfrak{S}(\mathcal{H})$ arose as a series rather than the product given above.

consider the moments

$$\frac{1}{N}\sum_{n\le N}(\pi(n+h)-\pi(n))^r = \frac{1}{N}\sum_{n\le N}\Bigg(\sum_{\substack{\ell=1\\ n+\ell \text{ prime}}}^{h} 1\Bigg)^r, \tag{6}$$

where r is a natural number. If the Cramér prediction is right, then we may expect these moments to be approximately

$$\frac{1}{N}\mathbb{E}\Bigg(\sum_{2\le n\le N}\Bigg(\sum_{\ell=1}^{h} X(n+\ell)\Bigg)^r\Bigg), \tag{7}$$

where $\mathbb{E}$ denotes expectation, and the $X(n)$'s are independent random variables as in Cramér's model. If these moments are roughly equal for $r \le R$ for any $R = R(N)$ tending to infinity, then we would know that $\pi(n+h)-\pi(n)$ has a Poisson distribution with parameter λ. This is because of a well-known principle from probability, that nice distributions including the Poisson distribution are determined by their moments.

Let us expand out the rth powers in (6) and (7). We then get numbers $\ell_1, \dots, \ell_r$ below h not necessarily distinct and would like to understand how often $n + \ell_1, \dots, n + \ell_r$ are all prime (for (6)), or to understand $\mathbb{E}(X(n + \ell_1)\cdots X(n + \ell_r))$ (for (7)). Let us suppose that there are exactly k distinct numbers among the $\ell_1, \dots, \ell_r$ and write these distinct numbers as $(1 \le)h_1 < h_2 < \cdots < h_k(\le h)$. The number of choices for $\ell_1, \dots, \ell_r$ that lead to the same ordered set of distinct numbers $h_1, \dots, h_k$ is the number of different ways of mapping $\{1, 2, \dots, r\}$ onto $\{1, \dots, k\}$; let us denote this[3] by $\sigma(r, k)$. Thus we see that (6) may be written as

$$\sum_{k=1}^{r}\sigma(r,k)\sum_{1\le h_1<h_2<\cdots<h_k\le h}\Bigg(\frac{1}{N}\sum_{\substack{n\le N\\ n+h_1,\dots,n+h_k \text{ prime}}} 1\Bigg), \tag{8}$$

while (7) may be written as

$$\sum_{k=1}^{r}\sigma(r,k)\sum_{1\le h_1<h_2<\cdots<h_k\le h}\Bigg(\frac{1}{N}\sum_{2\le n\le N}\mathbb{E}(X(n+h_1)\cdots X(n+h_k))\Bigg). \tag{9}$$

Since the same quantity appears in both (8) and (9) and is non-negative, we don't need to worry about what $\sigma(r, k)$ is.

[3] This is a 'Stirling number of the second kind.'

Invoking the Hardy–Littlewood conjecture (4)[4] we get that (8) is

$$\sim \sum_{k=1}^{r} \frac{\sigma(r,k)}{(\log N)^k} \sum_{1\le h_1<h_2<\cdots<h_k\le h} \mathfrak{S}(\{h_1,\ldots,h_k\}).$$

Clearly the quantity in (9) is

$$\sim \sum_{k=1}^{r} \frac{\sigma(r,k)}{(\log N)^k} \sum_{1\le h_1<h_2<\cdots<h_k\le h} 1.$$

Thus, to show that (6) and (7) are approximately equal, we need only show that

$$\sum_{1\le h_1<h_2<\cdots<h_k\le h} \mathfrak{S}(\{h_1,\ldots,h_k\}) \sim \sum_{1\le h_1<h_2<\cdots<h_k\le h} 1. \tag{10}$$

This is Gallagher's crucial result in (Gallagher, 1976). It shows that although the Hardy–Littlewood probabilities are different from the Cramér probabilities, on average they are roughly equal. This explains why the Cramér model makes accurate predictions for the distribution of primes in such short intervals.

EXERCISE 1.2. For a prime p put $\mathfrak{S}(\mathcal{H};p) = (1-\nu_{\mathcal{H}}(p)/p)(1-1/p)^{-k}$. Prove that as $h\to\infty$

$$\sum_{1\le h_1<h_2<\cdots<h_k\le h} \mathfrak{S}(\mathcal{H};p) \sim \sum_{1\le h_1<h_2<\cdots<h_k\le h} 1.$$

Explain why this morally implies (10); better still prove (10) rigorously (or read Gallagher's argument (Gallagher, 1976)).

EXERCISE 1.3. We have sketched how the Hardy-Littlewood conjecture implies that for a given positive real number λ, and a fixed non-negative integer k,

$$\frac{1}{N}\#\{n\le N : \pi(n+\lambda\log N)-\pi(n)=k\} \sim \frac{\lambda^k}{k!}e^{-\lambda}.$$

Deduce that

$$\frac{1}{N}\#\left\{2\le n\le N : \frac{p_{n+1}-p_n}{\log n}\in[\alpha,\beta]\right\} \sim \int_\alpha^\beta e^{-t}\,dt.$$

[4] Precisely, we need this conjecture uniformly for all $h_1,\ldots,h_k$ below h, and for all $k\le R$ with $R=R(N)$ tending slowly to infinity.

Proof of (10) *when* $k = 2$. From the definition (5) note that $\mathfrak{S}(\{h_1, h_2\}) = \mathfrak{S}(\{0, h_2 - h_1\})$ and so, letting $\ell = h_2 - h_1$ we see that the LHS of (10) is (in the case $k = 2$)

$$\sum_{\ell \le h} \mathfrak{S}(\{0,\ell\}) \Bigg(\sum_{\substack{1 \le h_1 < h_2 \le h \\ h_2 - h_1 = \ell}} 1 \Bigg) = \sum_{\ell \le h} \mathfrak{S}(\{0,\ell\})(h - \ell).$$

To evaluate the above asymptotically, it is useful to study the generating Dirichlet series

$$F(s) := \sum_{\ell=1}^{\infty} \frac{\mathfrak{S}(\{0,\ell\})}{\ell^s}.$$

The definition (5) gives that $\mathfrak{S}(\{0,\ell\}) = \prod_{p|\ell}(1 - 1/p)^{-1} \prod_{p\nmid\ell}(1 - 2/p)(1 - 1/p)^{-2}$. From this, we may see that $F(s)$ converges absolutely in the half-plane $\mathrm{Re}(s) > 1$, and moreover in that region has the Euler product

$$F(s) = \prod_p \Bigg(\Big(1 - \frac{2}{p}\Big)\Big(1 - \frac{1}{p}\Big)^{-2} + \frac{1}{p^s}\Big(1 - \frac{1}{p}\Big)^{-1} + \frac{1}{p^{2s}}\Big(1 - \frac{1}{p}\Big)^{-1} + \frac{1}{p^{3s}}\Big(1 - \frac{1}{p}\Big)^{-1} + \cdots \Bigg).$$

Multiplying and dividing by $\zeta(s) = \prod_p (1 - 1/p^s)^{-1}$ we see (with a little calculation) that in $\mathrm{Re}(s) > 1$,

$$F(s) = \zeta(s) \prod_p \Big(1 - \frac{1}{(p-1)^2} + \frac{1}{p^{s-1}(p-1)^2} \Big) = \zeta(s)G(s),$$

say. The Euler product for $G(s)$ converges absolutely in $\mathrm{Re}(s) > 0$ and so in that region we have obtained a meromorphic continuation of $F(s)$ with a simple pole at $s = 1$ coming from the simple pole of $\zeta(s)$ there. We now make use of the formula that for any $c > 0$

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{y^s}{s(s+1)}\, ds = \begin{cases} (1 - 1/y) & \text{if } y > 1 \\ 0 & \text{if } 0 < y \le 1. \end{cases}$$

This is easily proved by moving the line of integration to the left if $y > 1$ and to the right if $y \le 1$; the term $1 - 1/y$ when $y > 1$ arises from the residues of the poles at $s = 0$ and $s = -1$. Therefore, if $c > 1$, we see that

$$\begin{aligned} h \sum_{\ell \le h} \mathfrak{S}(\{0,\ell\}) \Big(1 - \frac{\ell}{h} \Big) &= h \sum_{\ell=1}^{\infty} \mathfrak{S}(\{0,\ell\}) \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \Big(\frac{h}{\ell}\Big)^s \frac{ds}{s(s+1)} \\ &= \frac{h}{2\pi i} \int_{c-i\infty}^{c+i\infty} \zeta(s)G(s) \frac{h^s}{s(s+1)}\, ds, \end{aligned}$$

where the interchange of summation and integration is justified by the absolute convergence of $F(s)$ in the region $\mathrm{Re}(s) > 1$. To evaluate the contour integral, we shift the line of integration to $\mathrm{Re}(s) = \epsilon > 0$. In the region traversed we encounter only a simple pole at $s = 1$ (because of $\zeta(s)$) and so our integral is

$$h \,\mathrm{Res}_{s=1}\left(\frac{\zeta(s)G(s)h^s}{s(s+1)}\right) + \frac{h}{2\pi i}\int_{\epsilon-i\infty}^{\epsilon+i\infty} \zeta(s)G(s)\frac{ds}{s(s+1)}.$$

Since $G(1)$ is easily seen to be 1, the residue above equals $G(1)h^2/2 = h^2/2$. By bounding $\zeta(s)$ and $G(s)$ on the line $\mathrm{Re}(s) = \epsilon$ we may estimate the remaining integral on that line; we omit the standard, but technical, details and merely note that this term is $O(h^{1+\epsilon})$. We conclude that

$$\sum_{\ell\le h}\mathfrak{S}(\{0,\ell\})(h-\ell) = \frac{h^2}{2} + O(h^{1+\epsilon}) = \sum_{h_1<h_2\le h} 1 + O(h^{1+\epsilon}).$$

This proves (10) in the case $k = 2$.

EXERCISE 1.4. Analyze $G(s)$ further by writing it as $\zeta(s+1)H(s)$, where $H(s)$ is now analytic in $\mathrm{Re}(s) > -\frac{1}{2}$. Evaluating residues, as above, prove that

$$\sum_{\ell\le h}\mathfrak{S}(\{0,\ell\})(h-\ell) = \frac{h^2}{2} - \frac{h\log h}{2} + \frac{Bh}{2} + O(h^{1/2+\epsilon}),$$

with $B = 1 - \gamma - \log 2\pi$; here γ is Euler's constant.

1.3. CONCLUDING REMARKS

Two important consequences of our predictions for the spacings between primes are that

$$\limsup_{n\to\infty}\frac{p_{n+1}-p_n}{\log n} = \infty, \quad \text{and} \quad \liminf_{n\to\infty}\frac{p_{n+1}-p_n}{\log n} = 0.$$

Happily both these results have now been proved. The first involves constructing long strings of composite numbers, and was first proved by Westzynthius with important refinements due to Erdős and Rankin. The second is a recent breakthrough of Goldston, Pintz and Yıldırım, see (Goldston et al., 2006). The reader may consult the survey by Heath-Brown (Heath-Brown, 1988) for the lim sup result and much else besides, and my survey (Soundararajan, 2006) for the lim inf result.

2. The Distribution of Primes in Longer Intervals

2.1. CRAMÉR'S PREDICTION

In the first lecture we considered the distribution of primes in intervals of length a constant times the average spacing. We now discuss what happens in longer intervals. Precisely, we consider $\pi(n+h)-\pi(n)$ for $n \le N$ and where $h/\log N$ is large, but h/N is small.

EXERCISE 2.1. Using Stirling's formula, show that as λ gets large, a Poisson distribution with parameter λ begins to look like a normal distribution with mean λ and variance λ.

Thus Cramér's model would suggest that, if $h/\log N$ is large but h/N is small, then for $n \le N$, $\pi(n+h)-\pi(n)$ has an approximately normal distribution with mean $\sim h/\log N$ and variance $\sim h/\log N$. Another way to arrive at this prediction is to calculate the moments (note that for most $n \le N$, $\sum_{\ell=1}^{h} 1/\log(n+\ell) \sim h/\log N$)

$$\frac{1}{N}\mathbb{E}\Bigg(\sum_{2\le n\le N}\bigg(\sum_{\ell=1}^{h} X(n+\ell)-\sum_{\ell=1}^{h}\frac{1}{\log(n+\ell)}\bigg)^{k}\Bigg), \tag{11}$$

which we claim is

$$= \frac{k!}{2^{k/2}(k/2)!}\bigg(\frac{h}{\log N}\bigg)^{k/2}\bigg(1+O_k\bigg(\frac{\log N}{h}\bigg)\bigg) \tag{12}$$

if k is even, and if k is odd it is

$$\ll \bigg(\frac{h}{\log N}\bigg)^{(k-1)/2}. \tag{13}$$

EXERCISE 2.2. Justify (11)–(13) by arguing as follows. For $n \ge 3$, set $X_0(n) = 1-1/\log n$ with probability $1/\log n$ and $X_0(n) = -1/\log n$ with probability $1-1/\log n$: that is, $X_0(n) = X(n)-1/\log n$. Note that $\mathbb{E}(X_0(n)) = 0$. Expand

$$\frac{1}{N}\mathbb{E}\Bigg(\sum_{2\le n\le N}\bigg(\sum_{1\le \ell\le h} X_0(n+\ell)\bigg)^{k}\Bigg) = \frac{1}{N}\sum_{\ell_1,\ldots,\ell_k\le h}\ \sum_{2\le n\le N}\mathbb{E}(X_0(n+\ell_1)\cdots X_0(n+\ell_k)).$$

The expectation above is zero if any of the ℓ_i's occurs only once among $\ell_1,\ldots,\ell_k$. When k is even there is a leading contribution from terms where the $\ell_1,\ldots,\ell_k$ contain $k/2$ distinct numbers each occurring twice.

2.2. CALCULATING THE VARIANCE VIA HARDY–LITTLEWOOD

However, we do not believe that this prediction, given by the Cramér model, is accurate. At this juncture, it is more convenient to deal with $\psi(n+h)-\psi(n)$, where $\psi(x) = \sum_{n\le x}\Lambda(n)$ with $\Lambda(n)$ denoting the von Mangoldt function. Note that the prime number theorem is equivalent to $\psi(x) \sim x$, and that the Hardy–Littlewood conjecture (4) may be recast as ($\mathcal{H} = \{h_1, \dots, h_k\}$ is a set of k distinct numbers)

$$\sum_{n\le x} \Lambda(n+h_1)\cdots\Lambda(n+h_k) \sim \mathfrak{S}(\mathcal{H})x. \tag{14}$$

The Cramér model predicts that $\psi(n+h)-\psi(n)$ is approximately normal with mean $\sim h$ and variance $\sim h\log N$.

To see the flaw in this prediction, let us now calculate the variance using the Hardy–Littlewood conjectures. Note that

$$\frac{1}{N}\sum_{n\le N}(\psi(n+h)-\psi(n)-h)^2 = \frac{1}{N}\sum_{n\le N}\Big(\sum_{\ell\le h}\Lambda(n+\ell)\Big)^2 - 2\frac{h}{N}\sum_{n\le N}\sum_{\ell\le h}\Lambda(n+\ell)+h^2.$$

The middle term in the RHS above is $-2h^2(\psi(N)+O(h\log N))/N \sim -2h^2$. As for the first term in the RHS we may square it out, and invoke the Hardy–Littlewood conjecture (14). If we forget all the error terms, then the above is

$$\frac{1}{N}\sum_{n\le N}\sum_{\ell\le h}\Lambda(n+\ell)^2 + 2\sum_{\ell\le h}\mathfrak{S}(\{0,\ell\})(h-\ell) - h^2.$$

The prime number theorem and partial summation gives that the first term above is $\sim h(\log N - 1)$, while from Exercise 1.4 we see that the second term above is $\sim h^2 - h\log h + Bh$. So, ignoring all error terms, we conclude that the variance satisfies

$$\frac{1}{N}\sum_{n\le N}(\psi(n+h)-\psi(n)-h)^2 \sim h\Big(\log\frac{N}{h}+B-1\Big), \tag{15}$$

which is different from the $\sim h\log N$ predicted by Cramér's model.

EXERCISE 2.3. Assume that the Hardy–Littlewood conjecture (14) holds in the quantitative form

$$\sum_{n\le x}\Lambda(n+h_1)\cdots\Lambda(n+h_k) = \mathfrak{S}(\mathcal{H})x + O(x^{1/2+\epsilon}),$$

uniformly for $k \le K$, and distinct h_j satisfying $1 \le h_j \le x$. Using this, obtain (15) with an error term of $O(h^{1/2+\epsilon}+h^2N^{-1/2+\epsilon}+h^3N^{-1})$. Thus, even assuming the quantitative Hardy–Littlewood conjectures, one knows (15) only for $h \le N^{1/2-\epsilon}$.

So although the Hardy–Littlewood probabilities and the Cramér probabilities are roughly equal on average, significant deviations show up when we consider h to be a small power of N. We believe that (15) is the right asymptotic for the variance and the Cramér model predicts the wrong answer.

2.3. THE VARIANCE AND ZEROS OF THE ZETA FUNCTION

Here is the sketch of a very different calculation which leads to the same answer as (15). Riemann's explicit formula (see (Davenport, 2000)) says that

$$\psi(x) = x - \sum_{\rho} \frac{x^{\rho}}{\rho} + \text{negligible terms.}$$

Here ρ runs over the non-trivial zeros of the Riemann zeta-function. We assume the Riemann hypothesis and write $\rho = \frac{1}{2} + i\gamma$. The sum over zeros is only conditionally convergent, but we will argue loosely omitting such considerations. It follows that

$$\psi(x+h) - \psi(x) - h = -\sum_{\rho} \frac{(x+h)^{\rho} - x^{\rho}}{\rho} + \text{negligible terms.}$$

The sum over zeros above is weighted down with a factor $1/\rho$, and so we may expect that large zeros make a minor contribution. It turns out that zeros with $|\rho| \geq x/h$ are not so important. For the small zeros, we replace $(x+h)^{\rho} - x^{\rho}$ by the Taylor approximation $\rho h x^{\rho-1}$. Therefore we may expect that

$$\begin{aligned} \frac{1}{X}\int_X^{2X} (\psi(x+h) - \psi(x) - h)^2 \, dx &\approx \frac{h^2}{X^2}\int_X^{2X} \Bigg| \sum_{|\gamma| \leq X/h} x^{i\gamma} \Bigg|^2 dx \\ &= \frac{h^2}{X} \sum_{|\gamma_1|,|\gamma_2| \leq X/h} X^{i(\gamma_1-\gamma_2)} \frac{2^{1+i(\gamma_1-\gamma_2)} - 1}{1 + i(\gamma_1 - \gamma_2)}. \end{aligned} \tag{16}$$

There are $\ll \log T$ zeros of the zeta-function with ordinates lying between T and $T+1$. Using this observation in (16), and estimating the magnitude of the sums over zeros there, we "deduce" that, assuming RH,[5]

$$\frac{1}{X}\int_X^{2X} (\psi(x+h) - \psi(x) - h)^2 \, dx \ll h(1 + \log X/h)^2. \tag{17}$$

A result like this was established by Selberg (Selberg, 1989). If we want an asymptotic in (16), then we need some understanding of the spacings

[5] In fact we would only deduce $\ll h(1 + \log X/h)^3$ but the extra "log" may be removed by smoothing.

$\gamma_1 - \gamma_2$ between zeros of the Riemann zeta-function. Such an understanding is provided by the pair correlation conjecture of Montgomery (Montgomery, 1973), which predicts that these ordinates are distributed like eigenvalues of large random matrices. Using such information Mueller obtained an asymptotic formula much like (15), and Goldston and Montgomery (Goldston and Montgomery, 1987) showed conversely that a formula like (15) also conveys information about the zeros of $\zeta(s)$. For more discussion on this set of ideas consult Goldston's recent survey (Goldston, 2005).

2.4. HIGHER MOMENTS

Recently, Montgomery and I (see (Montgomery and Soundararajan, 2002; Montgomery and Soundararajan, 2004)) used a quantitative form of the Hardy–Littlewood conjecture (see Exercise 2.3) to study higher moments of $\psi(n+h) - \psi(n) - h$. We now describe these results briefly. They support the conjecture that if $(\log N)^{1+\delta} \le h \le N^{1-\delta}$ then for $n \le N$ the distribution of $\psi(n+h) - \psi(n)$ is approximately normal with mean h and variance $h \log(N/h)$.

We assume that $(\log N)^{1+\delta} \le h \le N^{1-\delta}$ and wish to evaluate

$$\frac{1}{N} \sum_{n \le N} (\psi(n+h) - \psi(n) - h)^r. \tag{18}$$

For even r we expect that this is $\sim r!/(2^{r/2}(r/2)!)(h \log N/h)^{r/2}$, while for odd r we expect it to be $o((h \log N/h)^{r/2})$. If we simply expanded $(\psi(n+h) - \psi(n) - h)^r$ in powers of $(\psi(n+h) - \psi(n))$ and h (as we did in the case $r = 2$) then we would get many terms all of size h^r, and a careful cancellation of these and lower order terms is needed before we get to the actual delicate main term of size essentially $h^{r/2}$. To circumvent this, we define $\Lambda_0(n) = \Lambda(n) - 1$, in analogy with Exercise 2.2. This eliminates the unnecessary higher order terms at the outset, and simplifies calculations considerably. For other situations where this trick helps, see my paper with Granville (Granville and Soundararajan, 2006a) in this volume. Using this notation, and expanding (18) we want to understand

$$\sum_{\ell_1, \ldots, \ell_r \le h} \frac{1}{N} \sum_{n \le N} \Lambda_0(n + \ell_1) \cdots \Lambda_0(n + \ell_r). \tag{19}$$

EXERCISE 2.4. Define the modified singular series $\mathfrak{S}_0(\mathcal{H})$ by

$$\mathfrak{S}_0(\mathcal{H}) = \sum_{\mathcal{J} \subset \mathcal{H}} (-1)^{|\mathcal{H}| - |\mathcal{J}|} \mathfrak{S}(\mathcal{J}), \quad \text{so that } \mathfrak{S}(\mathcal{H}) = \sum_{\mathcal{J} \subset \mathcal{H}} \mathfrak{S}_0(\mathcal{J}).$$

Here we understand that $\mathfrak{S}(\emptyset) = \mathfrak{S}_0(\emptyset) = 1$. Show that the quantitative Hardy–Littlewood conjecture of Exercise 2.3 is the same as

$$\sum_{n\le x} \Lambda_0(n+h_1)\cdots\Lambda_0(n+h_k) = \mathfrak{S}_0(\mathcal{H})x + O(x^{1/2+\epsilon}),$$

keeping the hypotheses there.

For simplicity, consider first the terms in (19) when the ℓ_i are distinct. If we use the asymptotic of Exercise 2.4 we are led to the problem of evaluating

$$\sum_{\substack{h_1,\dots,h_k\le h\\ h_i \text{ distinct}}} \mathfrak{S}_0(\mathcal{H}),$$

which is a problem analogous to, but more delicate than, Gallagher's calculation (10). The main result in (Montgomery and Soundararajan, 2004) is the asymptotic

$$\sum_{\substack{h_1,\dots,h_k\le h\\ h_i \text{ distinct}}} \mathfrak{S}_0(\mathcal{H}) = \begin{cases} \{1+o(1)\}k!/(2^{k/2}(k/2)!)(-h\log h + B + 1)^{k/2} & \\ & \text{if } k \text{ is even} \\ o((h\log h)^{k/2}) & \text{if } k \text{ is odd.} \end{cases} \tag{20}$$

EXERCISE 2.5. Show the following refinement of Gallagher's (10):

$$\sum_{\substack{h_1,\dots,h_k\le h\\ h_i \text{ distinct}}} \mathfrak{S}(\mathcal{H}) = h^k - \binom{k}{2}h^{k-1}\log h + \binom{k}{2}Bh^{k-1} + O(h^{k-3/2+\epsilon}).$$

Returning to (19), we must analyze the terms when the ℓ_i are not necessarily distinct. Suppose that $h_1,\dots,h_k$ are the distinct elements among $\ell_1,\dots,\ell_r$ and that each h_i appears m_i (≥ 1) times among the ℓ_i. After a little combinatorics, we may write (19) as

$$\sum_{k=1}^{r} \sum_{\substack{m_1,\dots,m_k\ge 1\\ \sum m_i = r}} \binom{r}{m_1,\dots,m_k}\frac{1}{k!} \sum_{\substack{h_1,\dots,h_k\le h\\ h_j \text{ distinct}}} \frac{1}{N}\sum_{n=1}^{N}\prod_{i=1}^{k}\Lambda_0(n+h_i)^{m_i}. \tag{21}$$

We must distinguish the indices where $m_i = 1$ and the remaining indices where $m_i > 1$. Let $\mathcal{I}$ denote the subset of $\{1,\dots,k\}$ such that $m_i = 1$ for $i \in \mathcal{I}$. For $i \notin \mathcal{I}$ (so $m_i \ge 2$) we think of $\Lambda_0(n+h_i)^{m_i}$ as being essentially $(\log N)^{m_i-1}\Lambda(n+h_i)$: the point is that both quantities have about the same expected value $(\log N)^{m_i-1}$, unlike the case when $m_i = 1$ where the expected value of $\Lambda(n+h_i)$ and $\Lambda_0(n+h_i)$ are 1 and 0 respectively. Therefore the inner sum over n in (21) is essentially

$$\frac{(\log N)^{r-k}}{N} \sum_{n=1}^{N} \prod_{i\in\mathcal{I}} \Lambda_0(n+h_i) \prod_{\substack{1\le i\le k\\ i\notin\mathcal{I}}} (\Lambda_0(n+h_i)+1)$$
$$= \frac{(\log N)^{r-k}}{N} \sum_{\mathcal{I}\subset\mathcal{J}\subset\{1,\dots,k\}} \sum_{n=1}^{N} \prod_{j\in\mathcal{J}} \Lambda_0(n+h_j).$$

Now we invoke the Hardy–Littlewood conjecture of Exercise 2.4, and use (20).

EXERCISE 2.6. Complete the details in evaluating (19). Show that when r is odd or any of the m_i's is ≥ 3 we get a contribution of $o(h\log N)^{r/2}$. In the case r is even, the main term $r!/(2^{r/2}(r/2)!)(h\log N/h)^{r/2}$ arises from contributions to (21) where the m_i are all 1 or 2.

The proof of (20) is quite complicated, and we do not go into it here. Let us however point out one important ingredient. While motivating the Hardy–Littlewood conjecture in Lecture 1, we considered the toy problem of reduced residues (mod q). If $1 = a_1 < a_2 < \cdots < a_{\phi(q)} < q$ are the reduced residues below q, then we may ask for the distribution of $(a_{i+1}-a_i)(\phi(q)/q)$; we have multiplied by $\phi(q)/q$ so that this is 1 'on average.' If, for example, q is the product of the first ℓ primes, then as in Lecture 1 we may think of these a_i as being like primes, and expect that, for $0 < \alpha < \beta$,

$$\#\left\{1 \le i \le \phi(q)-1 : (a_{i+1}-a_i)\frac{\phi(q)}{q} \in [\alpha,\beta]\right\} \sim \phi(q)\int_{\alpha}^{\beta} e^{-x}\,dx.$$

A beautiful result of Hooley (Hooley, 1965) shows that this holds provided $\phi(q)/q$ is small. Obviously, some restriction on $\phi(q)/q$ is necessary; for example, if q is prime then clearly $a_{i+1}-a_i = 1$ for $1 \le i \le \phi(q)-1$. Moreover, Montgomery and Vaughan (Montgomery and Vaughan, 1986) have even estimated the moments:

$$M_k(q;h) = \sum_{n=1}^{q}\Biggl(\sum_{\substack{\ell\le h\\ (n+\ell,q)=1}} 1 - h\frac{\phi(q)}{q}\Biggr)^{k}.$$

The proof of (20) builds on the techniques developed there.

In our discussion above we have ignored error terms altogether. If one argues carefully using the quantitative Hardy-Littlewood conjectures of Exercises 2.3 and 2.4, we can evaluate the rth moment (18) provided that $h \leq N^{1/r-\epsilon}$. We expect that the same asymptotics hold even when h is larger with $h \leq N^{1-\epsilon}$. Thus these arguments suggest that for $(\log N)^{1+\delta} \leq h \leq N^{1-\delta}$, the distribution of $\psi(n+h) - \psi(n)$ (for $n \leq N$) is approximately normal with mean h and variance $h \log N/h$. For numerical support for this conjecture, see (Chan, 2002; Montgomery and Soundararajan, 2002). For other work related to this circle of ideas, see (Chan, 2002; Chan, 2006).

2.5. CONNECTIONS WITH ZEROS OF $\zeta(S)$?

We mentioned earlier the work of Goldston and Montgomery relating the variance of primes in short intervals to the pair correlation of zeros of $\zeta(s)$. Our calculations on the higher moments of primes in short intervals suggest that if $X \geq T^{1+\epsilon}$ then[6]

$$\int_X^{2X} \Big(\sum_{|\gamma| \leq T} x^{i\gamma} \Big)^k dx = \int_X^{2X} \Big(\sum_{0 \leq \gamma \leq T} 2\cos(\gamma \log x) \Big)^k dx$$

is $\sim k!/(2^{k/2}(k/2)!) X(2N(T))^{k/2}$ if k is even, and is $o(XN(T)^{k/2})$ if k is odd. Here $N(T) \sim (T/2\pi)\log T$ denotes the number of zeros of $\zeta(s)$ with $0 \leq \gamma \leq T$. Viewed this way, Montgomery's pair correlation conjecture may be thought of as saying that for $x \geq T^{1+\epsilon}$ the sum $\sum_{0 \leq \gamma \leq T} \cos(\gamma \log x)$ behaves like a sum of uncorrelated random variables. The higher moments suggest that it behaves in fact like a sum of independent random variables.[7] These statements are quite vague, and it would be nice to flesh out the precise connection between these higher moments and zeros of $\zeta(s)$. For other connections between zeros of $\zeta(s)$ and Hardy–Littlewood type conjectures see (Bogomolny and Keating, 1996).

2.6. CHEBYSHEV'S BIAS

We have considered above the distribution of primes in short intervals. What happens to the distribution in long intervals $[1, x]$? That is what can be said about the distribution of $\psi(x) - x$. Assuming RH, we get from Riemann's explicit formula that this is essentially $-2x^{1/2} \operatorname{Re} \sum_{0<\gamma} x^{i\gamma}/(1/2 + i\gamma)$. It is

[6] We take this opportunity to point out that the important constraint $X \geq T^{1+\epsilon}$ has been erroneously omitted in a similar discussion in (Montgomery and Soundararajan, 2004, p. 594).

[7] This is analogous to a result of E. Rains (Rains, 1997) in random matrix theory.

expected[8] that the zeros of $\zeta(s)$ are all simple, and have no non-trivial $\mathbb{Q}$-linear relations among them. In that case the sum over zeros above may be modeled by Re $\sum_{0<\gamma} X(\gamma)/(1/2+i\gamma)$ where the $X(\gamma)$ are independent random variables, taking uniformly distributed values on the unit circle. Precisely, as t varies from 1 to T, the distribution of $(\psi(e^t) - e^t)/(2e^{t/2})$ is like the distribution of our random sum above.[9] This is a certain non-universal distribution, which has been investigated in, for example, (Monach, 1980; Rubinstein and Sarnak, 1994). To gain a flavor of this distribution the reader may contemplate $\sum_{n=1}^{\infty} X_n/n$ where the X_n are independent random variables taking the values ± 1 with equal probability.

The distribution above is symmetric about the origin, and so $\psi(x)$ is as likely to be larger than x as it is to be smaller than x. However, $\psi(x) = \theta(x) + \theta(x^{1/2}) + \theta(x^{1/3}) + \cdots$ where $\theta(x) = \sum_{p\le x} \log p$. Thus it is much more likely for $\theta(x)$ to be smaller than x than for it to be larger than x. By partial summation one gets that $\pi(x) < \mathrm{li}(x)$ much more often than $\pi(x) > \mathrm{li}(x)$. In fact, in a certain sense the probability that $\pi(x)$ 'beats' $\mathrm{li}(x)$ is only $0.00000026\ldots$! We stop here, referring the reader to Rubinstein and Sarnak (Rubinstein and Sarnak, 1994), and the delightful recent survey (Granville and Martin, 2006) for more information.

To summarize, we found three distinct behaviors for the distribution of primes in intervals. At the "microscopic" scale ($h \asymp \log N$) there is Poisson behavior, at the "mesoscopic" scale ($h/\log N \to \infty$, $h = o(N)$) there is Gaussian (normal) behavior, and at the "macroscopic" scale ($h \gg N$) there is a specific non-universal distribution law. Such division into three regimes occurs in many other problems as well; for example, in the distribution of lattice points in the plane. As a starting point, we refer the reader to the recent paper of Hughes and Rudnick (Hughes and Rudnick, 2004) and to the references therein.

3. Maier's Method and an "Uncertainty Principle"

If the Riemann Hypothesis is true, then from Selberg's result (17) we easily deduce that (for $h \le N$) the number of $n \le N$ with $|\psi(n+h) - \psi(n) - h| \ge \sqrt{h}(\log N)^{1+\delta}$ is $\ll N/(\log N)^{2\delta}$. It follows that if $N \ge h \ge (\log N)^{2+\delta}$ then "almost all" intervals $(n, n+h]$ with $n \le N$ contain about the correct number of primes, $\sim h/\log N$. If (15) holds then we can even conclude that if $h \ge (\log N)^{1+\delta}$ then "almost all" intervals $(n, n+h]$ with $n \le N$ contain

[8] There is perhaps no good reason for this belief, except that the contrary situation is harder to imagine!

[9] The change of variable $x = e^t$ means that $x^{i\gamma} = e^{it\gamma}$ now takes values uniformly on the unit circle as t varies.

approximately the correct number of primes. In Cramér's model, one can show that almost surely $\sum_{x \le n \le x+h} X(n) \sim h/\log x$ if $(\log x)^{2+\delta} \le h \le x$. Thus it seems quite plausible that if x is large and $x \ge h \ge (\log x)^{2+\delta}$ then $\psi(x+h) - \psi(x) \sim h$.

The classical prime number theorem with error term $x \exp(-C\sqrt{\log x})$ tells us that such a result holds if $h \ge x \exp(-C\sqrt{\log x})$. An important advance was made by Hoheisel who showed that the asymptotic $\psi(x+h) - \psi(x) \sim h$ holds if $x \ge h \ge x^\theta$ for some number $\theta < 1$. He was able to take $\theta = 1 - 1/33000$, but this has been improved subsequently, with the best result known, due to Huxley, being $\theta = 7/12 + \epsilon$. If the Riemann hypothesis is true then θ may be taken as $\frac{1}{2} + \epsilon$. The arguments pioneered by Hoheisel depend on the fact that while we don't know RH, we do know that most zeros of $\zeta(s)$ lie close to the $\frac{1}{2}$ line. For a nice account of these results see Heath-Brown (Heath-Brown, 1988). If the asymptotics given for (18) are true then we may take θ to be any positive number.

Thus it seems the conjecture that $\psi(x+h)-\psi(x) \sim h$ for $x \ge h \ge (\log x)^{2+\delta}$, if true, lies quite deep. This conjecture was widely believed until the mid 1980s when Maier (Maier, 1985) shattered this belief by showing that for any $A > 1$ there are arbitrarily large x such that the interval $(x, x + (\log x)^A]$ contains significantly more primes than usual (that is, $\ge (1 + \delta_A)(\log x)^{A-1}$ primes for some $\delta_A > 0$) and also intervals $(x, x + (\log x)^A]$ containing significantly fewer primes than usual. In this lecture we will sketch Maier's ingenious method, and describe some extensions of his idea. The reader may also consult (Granville, 1995) for another exposition of related ideas.

3.1. MAIER'S "MATRIX" METHOD

Let x be large, and h be on the scale of a power of $\log x$. Let P be an integer which we will eventually take to be the product of many small (below $\log x$) primes. Consider the $[x/P]$-by-h "matrix" whose (i, j)th entry is the number $([x/P] + i)P + j$. Thus the entries of this matrix are numbers lying between x and $2x + h$. Note that each row of the matrix consists of an interval $([x/P] + i)P$ to $([x/P] + i)P + h$. Each column of the matrix consists of an arithmetic progression with common difference P: namely, $x \le n \le 2x + h$ with $n \equiv j \pmod P$. The idea is to count the number of primes in this matrix in two ways: by counting the primes row by row, and by counting the primes column by column, and then comparing the two answers. If we assume that the asymptotic formula for primes in short intervals holds then we get an answer for the row by row calculation. The prime number theorem for arithmetic progressions allows us to do the column by column calculation. Of course the two answers should match. However when h is very small, like a power of $\log x$, there are choices of P for which the answers don't match! This leads to Maier's result.

Consider the row by row calculation. The number of primes is

$$\sum_{x/P\le n\le 2x/P} (\pi(nP+h)-\pi(nP)), \tag{22}$$

and if we assume that intervals of length h contain the right number of primes, this is

$$\sim \frac{x}{P}\frac{h}{\log x}. \tag{23}$$

Consider next the column by column calculation. If the progression $n \equiv j \pmod P$ is to contain primes, we must have $(j, P) = 1$. In that case the prime number theorem in arithmetic progressions would say that such a progression contains a proportion $1/\phi(P)$ of all primes. Of course, in order to use the prime number theorem in arithmetic progressions rigorously we must pay attention to the size of the modulus P compared with x. Assuming that this is not an issue, we find that the column by column contribution is

$$\sum_{\substack{j\le h\\ (j,P)=1}} (\pi(2x+h;P,j)-\pi(x;P,j)) \sim \frac{x}{\phi(P)\log x}\sum_{\substack{j\le h\\ (j,P)=1}} 1. \tag{24}$$

If we compare (24) and (23) we find the relation

$$\sum_{\substack{j\le h\\ (j,P)=1}} 1 \sim h\frac{\phi(P)}{P} \tag{25}$$

should hold. At first glance, (25) is eminently reasonable: the probability that j is coprime to P is $\phi(P)/P$. It is even easy to make this precise: write the condition $(j, P) = 1$ as $\sum_{\ell|(j,P)}\mu(\ell)$ and we easily get

$$\sum_{\substack{j\le h\\ (j,P)=1}} 1 = h\frac{\phi(P)}{P} + O(d(P)big), \tag{26}$$

where $d(P)$ is the number of divisors of P. Thus, if h is just a bit larger than $d(P)$ (which is always quite small, that is $\ll P^{\epsilon}$) then (25) will hold. So where is the contradiction? The point is that in Maier's application h is very small compared with P, and so (26) is useless.

For the purpose of illustration suppose that P is the product of all primes between $(\log x)^{9/10}$ and $(\log x)/100$. Then, by the prime number theorem, P is about size $x^{1/100+o(1)}$. For such moduli P we don't know the prime number theorem in arithmetic progressions used in (24), but such a result does hold if the Riemann hypothesis for Dirichlet L-functions is true; let us postpone a discussion of this point. Suppose now that h is a number of size $(\log x)^\theta$ with $2 < \theta < 2.7$. By inclusion-exclusion, the LHS of (25) is

$$\begin{aligned}&\sim\ h - \sum_{(\log x)^{0.9} \le p \le (\log x)/100} \frac{h}{p} + \sum_{(\log x)^{0.9} \le p < q \le (\log x)/100} \frac{h}{pq}\\ &\sim\ h\left(1 - \log\frac{10}{9} + \frac{1}{2}\left(\log\frac{10}{9}\right)^2\right),\end{aligned}$$

where we have used the prime number theorem to evaluate $\sum 1/p$ for p between $(\log x)^{9/10}$ and $(\log x)/100$. On the other hand, by Mertens' theorem, the RHS of (25) is

$$\sim h \prod_{(\log x)^{9/10} \le p \le (\log x)/100} \left(1 - \frac{1}{p}\right)^{-1} \sim \frac{9}{10}h.$$

The formula for the LHS has the first three terms in the usual expansion of $9/10 = e^{-\log(10/9)}$, so the two answers are certainly close, but obviously they are not equal! Indeed the LHS is a little bit larger.

EXERCISE 3.1. Conclude from the above that for any $2 < \theta < 2.7$ there exist arbitrarily large x such that the interval $[x, x + (\log x)^\theta]$ contains significantly more primes than expected. Taking such an interval and cutting it up into smaller intervals, deduce that the same conclusion holds for all $1 < \theta < 2.7$. Using the same P as above, and taking four terms in the inclusion-exclusion formula, show that if $\theta < 3.6$ there exist intervals $[x, x + (\log x)^\theta]$ with significantly fewer primes than expected. In this manner one can proceed for $\theta < 8.1$, just using inclusion-exclusion and easy calculations. Now replace $(\log x)^{0.9}$ in the definition of P with $(\log x)^{1-\delta}$ and prove Maier's theorem.

3.2. MORE ON THE CONTRADICTION

Now let us describe a different way of seeing a contradiction to (25). This method is very flexible, and works for many choices of P, and also generalizes readily. Let y be a large parameter; in the application we may think of y as being some power of $\log x$. From each dyadic block $[2^{-j}y, 2^{-j+1}y]$ with $j \le [\log y/(2\log 2)]$ select about half the primes. Take P to be the product of these selected primes. Thus P is composed of about half the primes in $[\sqrt{y}, y]$, and there are plenty of choices for P. Let $u \ge 1$ be a real number, set $h = y^u$ and consider whether (25) can hold. We will show that for arbitrarily large u

the LHS is appreciably larger than the RHS, and for arbitrarily large u it is smaller.

To see this we consider the Dirichlet series $\zeta_P(s) = \sum_{(n,P)=1} n^{-s}$. Plainly we have

$$\zeta_P(s) = \zeta(s)\prod_{p|P}\Big(1 - \frac{1}{p^s}\Big), \tag{27}$$

so that $\zeta_P(s)$ extends to a meromorphic function in all of $\mathbb{C}$ with a simple pole at $s = 1$. The point is that if something like (25) holds then $\zeta_P(s)$ must approximately look like $\zeta(s)\phi(P)/P$, and by choosing s appropriately we can obtain a contradiction to (27). More precisely, set

$$E(u) = \frac{1}{y^u}\bigg(\sum_{\substack{n\le y^u\\(n,P)=1}} 1 - [y^u]\frac{\phi(P)}{P}\bigg).$$

Then, for $\mathrm{Re}(s) > 1$,

$$\zeta_P(s) - \zeta(s)\frac{\phi(P)}{P} = \int_{1^-}^{\infty}\frac{1}{z^s}d\bigg(\sum_{\substack{n\le z\\(n,P)=1}} 1 - \sum_{n\le z}\frac{\phi(P)}{P}\bigg) = \int_1^{\infty}\frac{s}{z^s}E\Big(\frac{\log z}{\log y}\Big)dz,$$

upon integrating by parts. Changing variables $u = \log z/\log y$ we obtain that

$$\zeta_P(s) = \zeta(s)\frac{\phi(P)}{P} + s\log y\int_0^{\infty} E(u)y^{-u(s-1)}\,du. \tag{28}$$

To start with, (28) is valid for $\mathrm{Re}(s) > 1$, but since $E(u) \ll d(P)y^{-u}$ by (26), we see that (28) makes sense for $\mathrm{Re}\, s > 0$.

EXERCISE 3.2. Let $(\log y)/2 \ge \xi \ge 1$ be a real number, and take $s = 1 - \xi/\log y + i\pi/\log y$. Using (27) prove that

$$|\zeta_P(s)| \gg \frac{\log y}{\xi}\exp\bigg(\frac{e^{\xi}}{2\xi} + O\Big(\frac{e^{\xi}}{\xi^2}\Big)\bigg).$$

Then using (28) deduce that

$$\int_0^{\infty}|E(u)|e^{\xi u}\,du \gg \frac{1}{\xi}\exp\bigg(\frac{e^{\xi}}{2\xi} + O\Big(\frac{e^{\xi}}{\xi^2}\Big)\bigg).$$

Show that $\int_0^{\infty} E(u)e^{\xi u}du \ll 1/\xi$, so that in the LHS above both positive and negative values of $E(u)$ make roughly equal contributions.

Morally, Exercise 3.2 shows that $E(u)$ cannot be too small for large u. To make this precise, one also needs an upper bound for $E(u)$ so as to be able to bound the tail of the integral in Exercise 3.2. Developing this argument carefully, one may show that there is a positive constant A such that every interval $[u(1 - A/\log u), u(1 + A/\log u)]$ contains points $u_\pm$ satisfying

$$E(u_+) \geq \exp\Big(- u_+(\log u_+ + \log\log u_+ + O(1))\Big),$$

and

$$E(u_-) \leq - \exp\Big(- u_-(\log u_- + \log\log u_- + O(1))\Big).$$

For more details, see (Granville and Soundararajan, 2006b, Section 3, especially Corollary 3.3).

Earlier, we postponed discussion of the prime number theorem in arithmetic progressions. We refer the reader to Davenport (Davenport, 2000) for an account of this. In Chapter 20 there one finds Page's result that $\psi(x; q, a) \sim x/\phi(q)$ for all $q \leq \exp(C\sqrt{\log x})$ with the possible exception of multiples of a particular modulus q_1 which may depend on x. If we choose y a little less than $\sqrt{\log x}$ then our moduli P above are below $\exp(C\sqrt{\log x})$ and certainly we can find P that are not multiples of the exceptional modulus q_1. Thus our appeal to the prime number theorem in arithmetic progressions can be made rigorous.

The flexibility in choosing P is quite useful. Exploiting this, Granville and I (see (Granville and Soundararajan, 2006b)) showed that the asymptotic

$$\psi(x + h) - \psi(x) = h + O(h^{\frac{1}{2}+\epsilon}), \tag{29}$$

suggested by Cramér's model, sometimes fails to hold if $h \leq \exp\left((\log x)^{\frac{1}{2}-\epsilon}\right)$. This improves work of Hildebrand and Maier (Hildebrand and Maier, 1989) who had obtained this result assuming the Generalized Riemann Hypothesis, and a weaker result unconditionally. It seems safe to conjecture that (29) holds if $h \geq x^\delta$, and perhaps it holds when $h \geq \exp\left((\log x)^{\frac{1}{2}+\delta}\right)$.

3.3. AN UNCERTAINTY PRINCIPLE

Maier's method can be adapted to establish limitations to the equidistribution of primes in arithmetic progressions. For example, Friedlander and Granville (Friedlander and Granville, 1989) proved that for every $A \geq 1$ there exist large x and an arithmetic progression $a \pmod q$ with $(a, q) = 1$ and $q \leq x/(\log x)^A$ such that

$$\left|\pi(x; q, a) - \frac{\pi(x)}{\phi(q)}\right| \gg_A \frac{\pi(x)}{\phi(q)}.$$

More recently, Balog and Wooley (Balog and Wooley, 2000) showed that the sequence of integers which may be written as the sum of two squares also exhibits "Maier type" irregularities in intervals $(x, x + (\log x)^A)$ for any fixed, positive A. Previously Maier's work had seemed inextricably linked to the mysteries of primes, but Balog and Wooley's result suggests that such results should be part of a more general phenomenon. This has been formalized by Granville and me as an "uncertainty principle" for arithmetic sequences. What Maier's argument shows is that the primes cannot be simultaneously well distributed in short intervals, and in arithmetic progressions. Then a suitable version of the prime number theorem in arithmetic progressions is used to remove the second possibility, leaving us with the irregularities of distribution in short intervals. The first conclusion of irregularities in short intervals or progressions turns out to be a general feature of many interesting arithmetical sequences.

A rough description of this result is as follows: Let $\mathcal{A}$ denote a sequence $a(n)$ of non-negative real numbers, and let $\mathcal{A}(x) = \sum_{n\le x} a(n)$. If $\mathcal{A}$ is well-distributed in short intervals, then we may expect that

$$\mathcal{A}(x+y) - \mathcal{A}(x) \approx y\frac{\mathcal{A}(x)}{x}. \tag{30}$$

To understand the distribution of $a(n)$ in arithmetic progressions we begin with n that are multiples of a given number d. We suppose that there is a non-negative multiplicative function h such that

$$\mathcal{A}_d(x) = \sum_{\substack{n\le x\\ d|n}} a(n) \approx \frac{h(d)}{d}\mathcal{A}(x). \tag{31}$$

We assume that the asymptotic behavior of

$$\mathcal{A}(x; q, a) := \sum_{\substack{n\le x\\ n\equiv a \pmod q}} a(n)$$

depends only on the g.c.d. of a and q. Then (31) leads to the prediction that

$$\mathcal{A}(x; q, a) \approx \frac{f_q(a)}{q\gamma_q}\mathcal{A}(x), \tag{32}$$

with $\gamma_q = \prod_{p|q}(p-1)/(p-h(p))$ and $f_q(a)$ is a certain non-negative multiplicative function of a, defined in terms of h, such that $f_q(a) = f_q((a, q))$ so that $f_q(a)$ is periodic $\pmod q$. We can be flexible in how we want to assume (32); for example, sometimes it is convenient to assume it only for q that are coprime to a certain fixed set of primes.

To illustrate the framework consider the following examples.

EXAMPLE 3.3. Take $a(n) = 1$ for all n. It is natural to take $h(d) = 1$ for all d, $\gamma_q = 1$, and $f_q(a) = 1$. Then (31) and (32) are both good approximations with errors at most 1.

EXAMPLE 3.4. Take $a(n)$ to be the indicator function of the primes. Then $h(1) = 1$ and $h(d) = 0$ for $d > 1$. One has $\gamma_q = \phi(q)/q$ and $f_q(a) = 1$ if $(a, q) = 1$ and 0 otherwise. The prime number theorem in arithmetic progressions gives (32) for small values of q. The result of Friedlander and Granville places restrictions on the approximation (32) when q is large. Maier's results place restrictions on (31) for small y.

EXAMPLE 3.5. Take $a(n)$ to be the indicator function of the sums of two squares. The multiplicative function h is defined by $h(p^k) = 1$ if $p^k \equiv 1 \pmod 4$ and $h(p^k) = 1/p$ if $p^k \equiv 3 \pmod 4$. Here Balog and Wooley's result places restrictions on (31).

The main results of (Granville and Soundararajan, 2006b) give that if $h(p)$ is not always close to 1 (as in the regular Example 3.3) then there will be moduli q for which (32) cannot hold. Typically these moduli will be large as in the Friedlander–Granville result for primes in progressions. Furthermore, either there exist values y larger than an arbitrary power of $\log x$ for which (31) is false, or there exist small moduli q (below $\exp((\log x)^\delta)$) for which (32) is false. These results include the previous results on primes and sums of two squares, and also cover many other examples.

Consider sets containing roughly half of the prime numbers. There are uncountably many such sets, and so maybe we can find a set which is very well distributed in arithmetic progressions. One amusing example from (Granville and Soundararajan, 2006b) shows that this cannot be done, and the Friedlander– Granville limitations persist for any such set.

We content ourselves with this vague description of the uncertainty principle, referring the reader to (Granville and Soundararajan, 2006b) for more examples and a precise description of the results.

Acknowledgements

The author is partially supported by the National Science Foundation.

References

Balog, A. and Wooley, T. D. (2000) Sums of two squares in short intervals, *Canad. J. Math.* **52**, 673–694.

Bogomolny, E. B. and Keating, J. P. (1996) Random matrix theory and the Riemann zeros. II. n-point correlations, *Nonlinearity* **9**, 911–935.

Chan, T. H. (2002) Pair correlation and distribution of prime numbers, Ph.D. thesis, University of Michigan.

Chan, T. H. (2006) A note on primes in short intervals, *Int. J. Number Theory* **2**, 105–110.

Cramér, H. (1936) On the order of magnitude of the difference between consecutive prime numbers, *Acta Arith.* **2**, 23–46.

Davenport, H. (2000) *Multiplicative number theory*, Vol. 74 of *Grad. Texts in Math.*, New York, Springer.

Feller, W. (1966) *An introduction to probability theory and its applications*, New York–London–Sydney, Wiley.

Friedlander, J. and Granville, A. (1989) Limitations to the equi-distribution of primes. I, *Annals of Math. (2)* **129**, 363–382.

Gallagher, P. X. (1976) On the distribution of primes in short intervals, *Mathematika* **23**, 4–9.

Goldston, D. (2005) Notes on pair correlation of zeros and prime numbers, In *Recent perspectives in random matrix theory and number theory*, Vol. 322 of *London Math. Soc. Lecture Notes Ser.*, Cambridge, Cambridge Univ. Press, pp. 79–110.

Goldston, D. and Montgomery, H. L. (1987) On pair correlations of zeros and primes in short intervals, In *Analytic number theory and Diophantine problems*, Vol. 70 of *Prog. Math.*, Stillwater, OK, 1984, pp. 183–203, Boston, Birkhäuser.

Goldston, D., Pintz, J., and Yıldırım, C. (2006) Primes in tuples. I, *Ann. of Math. (2)*, to appear; preprint available at www.arxiv.org.

Granville, A. (1995) Unexpected irregularities in the distribution of prime numbers, In *Proceedings of the International Congress of Mathematicians. Vol. 1, 2*, Zürich, 1994, pp. 388–399, Basel, Birkhäuser.

Granville, A. and Martin, G. (2006) Prime number races, *Amer. Math. Monthly* **113**, 1–33.

Granville, A. and Soundararajan, K. (2006a) Sieving and the Erdős–Kac theorem, this book.

Granville, A. and Soundararajan, K. (2006b) An uncertainty principle for arithmetic sequences, *Ann. of Math.(2)*, to appear; preprint available at www.arxiv.org.

Hardy, G. H. and Littlewood, J. E. (1922) Some problems of Paritio Numerorum. III. On the expression of a number as a sum of primes, *Acta Math.* **44**, 1–70.

Heath-Brown, D. R. (1988) Differences between consecutive primes, *Jahresber. Deutsch. Math.-Verein.* **90**, 71–89.

Hildebrand, A. and Maier, H. (1989) Irregularities in the distribution of primes in short intervals, *J. Reine Angew. Math.* **397**, 162–193.

Hooley, C. (1965) On the difference between consecutive numbers prime to n. II, *Publ. Math. Debrecen* **12**, 39–49.

Hughes, C. P. and Rudnick, Z. (2004) On the distribution of lattice points in thin annuli, *Int. Math. Res. Not.* **2004**, 637–658.

Maier, H. (1985) Primes in short intervals, *Michigan Math. J.* **32**, 221–225.

Monach, W. (1980) Numerical investigation of several problems in number theory, Ph.D. thesis, University of Michigan.

Montgomery, H. L. (1973) The pair corelation of zeros of the zeta function, In *Analytic Number Theory*, Vol. 24 of *Proc. Sympos. Pure Math.*, St. Louis Univ., 1972, pp. 181–193, Providence, RI, Amer. Math. Soc.

Montgomery, H. L. and Soundararajan, K. (2002) Beyond pair correlation, In *Paul Erdős and his mathematics. I*, Vol. 11 of *Bolyai Soc. Math. Stud.*, Budapest, 1999, pp. 507–514, Budapest, János Bolyai Math. Soc.

Montgomery, H. L. and Soundararajan, K. (2004) Primes in short intervals, *Comm. Math. Phys.* **252**, 589–617.

Montgomery, H. L. and Vaughan, R. C. (1986) On the distribution of reduced residues, *Ann. of Math. (2)* **123**, 311–333.

Rains, E. M. (1997) High powers of random elements of compact Lie groups, *Probab. Theory Related Fields* **107**, 219–241.

Rubinstein, M. and Sarnak, P. (1994) Chebyshev's bias, *Experimental Math.* **3**, 173–197.

Selberg, A. (1989) On the normal density of primes in short intervals, and the difference between consecutive primes, In *Collected papers. Vol. I*, Berlin, Springer, pp. 160–178.

Soundararajan, K. (2006) Small gaps between prime numbers: the work of Goldston–Pintz–Yıldırım, *Bull. Amer. Math. Soc.*, to appear; preprint available at www.arxiv.org.

TORSION POINTS ON CURVES

Andrew Granville
Université de Montréal

Zeév Rudnick
Tel-Aviv University

1. Introduction

One of the themes of the summer school is the distribution of "special points" on varieties. In Heath-Brown's lectures we study rational points on projective hyper-surfaces; in Ullmo's course we study Galois orbits and Duke's lectures deal with CM-points on the modular curve. This lecture concerns one of the earliest examples, namely torsion points on group varieties.

DEFINITION 1.1. For a group A, the torsion points are

$$\mathrm{Tor}(A) = \{x \in A : x^n = 1 \text{ for some } n \geq 1\}$$

(we write the group law as multiplication).

If A is abelian then $\mathrm{Tor}(A)$ is a subgroup of A.

EXAMPLES.

(i) The multiplicative group $A = \mathbf{G}_m$ is the algebraic group whose points over a field are the nonzero elements of the field. Then for any field K, $\mathrm{Tor}\,\mathbf{G}_m(K)$ are the roots of unity contained in K.

(ii) $A = \mathbf{G}_m \times \mathbf{G}_m$ then $\mathrm{Tor}(A) = \mathrm{Tor}(\mathbf{G}_m) \times \mathrm{Tor}(\mathbf{G}_m) = \{(x, y) : x, y \in K$ are roots of unity$\}$.

(iii) Let A be an elliptic curve. Over the complex numbers we can uniformize A as $A = \mathbb{C}/L$ where L is a lattice. Then $\mathrm{Tor}(A(\mathbb{C})) = \mathbb{Q} \otimes L/L$.

More generally we can study *division points*:

DEFINITION 1.2. If $\Gamma \subset A$ is a finitely generated group, let

$$\mathrm{Tor}(A, \Gamma) = \{x \in A : x^n \in \Gamma \text{ for some } n \neq 0\}$$

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 85–92.

Thus $\mathrm{Tor}(A, \{1\}) = \mathrm{Tor}(A)$ are the torsion points of A.

Motivated by Mordell's conjecture, Lang (Lang, 1965) made the following

CONJECTURE A. *If V is an irreducible curve on an abelian group variety (e.g., $A = (\mathbf{G}_m)^n$ or an abelian variety) and $\Gamma \subset A$ is a finitely generated subgroup such that* $\mathrm{Tor}(A, \Gamma) \cap V$ *is infinite, then V is a translate of a subgroup of A by a division point.*

See Ullmo's lectures (Ullmo, 2006) for the statement of the Manin–Mumford conjecture, which generalizes this statement, and the survey (Tzermias, 2000) for more background.

The first instance of Lang's conjecture is for torsion points on $(\mathbf{G}_m)^r$, which turns out to be quite elementary. We will present two proofs of Lang's conjecture for that case.

2. A Proof Using Galois Theory

The first proof is that which appears in the original paper by Lang (Lang, 1965) where it is attributed to Ihara, Serre and Tate. The result is

THEOREM 2.1. *Let $V/\mathbb{C}$ be an irreducible curve in $A = \mathbf{G}_m \times \mathbf{G}_m$. If V contains infinitely many torsion points then V is a translate of a subgroup of $A = \mathbf{G}_m \times \mathbf{G}_m$ by a torsion point, i.e.,*

$$V = \{(x, y) : x^r = \zeta y^s\}$$

for some root of unity ζ.

To highlight the ideas we will only consider a special case: $V \subset \mathbf{G}_m \times \mathbf{G}_m$ is a rational curve of the forms $\{(f(t), g(t))\}$ where f and g are polynomials, which for added simplicity we assume to have rational coefficients: $f, g \in \mathbb{Q}[t]$. Then

$$V \cap \mathrm{Tor}(A) = \{(f(t), g(t)) \text{ are both roots of unity}\}$$

The subgroups of $\mathbf{G}_m \times \mathbf{G}_m$ are $\{(x, y) : x^r = y^s\}$ for some integers r, s. So we need to show

THEOREM 2.2. *Let $f, g \in \mathbb{Q}[t]$ be polynomials. If there are infinitely many values of t for which both $f(t)$ and $g(t)$ are roots of unity then there are nonzero integers $r, s \neq 0$ so that $f^r = g^s$.*

Proof. We assume there are infinitely many t so that both $f(t)$, $g(t)$ are roots of unity and want to force the relation $f^r = g^s$.

Take $n \gg 1$ so that there is some z_1 with

$$f(z_1) = \zeta_n^\alpha, \quad g(z_1) = \zeta_n^\beta$$

where ζ_n denotes a primitive nth root of unity and that this is the minimal way of writing such an expression, that is $\gcd(n, \alpha, \beta) = 1$ (exercise). Note that $z_1 \in \overline{\mathbb{Q}}$ is algebraic. Then we have a relation

$$f(z_1)^\beta = g(z_1)^\alpha$$

(and both sides equal $\zeta_n^{\alpha\beta}$), but this relation holds for only *one* point z_1 and we want it to hold for *all* points z.

Now apply the Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, which acts *transitively* on the primitive n-th roots of unity (see Ullmo's lectures (Ullmo, 2006)). Hence if σ_j is a Galois automorphism so that $\sigma_j(\zeta_n) = \zeta_n^j$, $\gcd(j, n) = 1$ and $z_j := \sigma_j(z_1)$ then because we assume f, g have rational coefficients we get

$$\sigma_j(f(z_1)) = f(\sigma_j(z_1)) = f(z_j), \quad \sigma_j(g(z_1)) = g(z_j)$$

and so

$$f(z_j)^\beta = \sigma_j(f(z_1)) = \zeta_n^{j\alpha\beta} = g(z_j)^\alpha.$$

Now we have the relation $f^\beta = g^\alpha$ holding for $\phi(n)$ distinct points[1] rather than just one point (exercise: why are the points z_j distinct?). However we still need it to hold for *all* z.

Consider the polynomial

$$F(t) = f(t)^\beta - g(t)^\alpha.$$

It has $\phi(n)$ distinct roots so if $\deg F < \phi(n)$ then we would have $F \equiv 0$ as required. Now if $F \neq 0$ then

$$\deg F = \max(\beta \deg f, \alpha \deg g)$$

can be as large as const$\cdot n$. which is still (slightly) too big relative to $\phi(n)$.

The remedy is to raise the relation $f(z_j)^\beta = g(z_j)^\alpha = \zeta_n^{\alpha\beta}$ to an m-th power:

$$f(z_j)^{m\beta} = g(z_j)^{m\alpha}$$

(both sides equal $\zeta_n^{m\alpha\beta}$). We get a new polynomial $f^{m\beta} - g^{m\alpha}$ with $\phi(n)$ distinct roots; it looks like we raised the degree which is certainly useless! However,

[1] $\phi(n)$ is the number of residues coprime to n

since $f(z_j)$, $g(z_j)$ are nth roots of unity, we have $f(z_j)^n = 1 = g(z_j)^n$ and if we substitute

$$m\beta \equiv r \bmod n, \quad m\alpha \equiv s \bmod n$$

with $|r|, |s| \le n/2$ then we find $f(z_j)^r = g(z_j)^s$ for all j coprime to n. This is still not useful as we have just showed that $\deg F \le \max(\deg f, \deg g)n/2$ instead of showing that $\deg F < \phi(n)$. However we will be done if we can show that there is some $m \ge 1$ so that the residues $(m\beta, m\alpha) \bmod n$ are both small! This is given by the following

EXERCISE. Given a primitive vector $(\alpha, \beta) \in (\mathbb{Z}/n\mathbb{Z})^2$, that is $\gcd(\alpha, \beta, n) = 1$, there is some $1 \le m \le n$ so that both residues $m\alpha \bmod n$ and $m\beta \bmod n$ are at most $n^{2/3}$ (and are different than $(0, 0) \bmod n$).

See Venkatesh's lecture (Venkatesh, 2006) and (Strombbergsson and Venkatesh, 2005) where it is shown that typically the size of both residues is about $\sqrt{n}$.

Consequently we find a relation $f(z_j)^r = g(z_j)^s$ with $|r|, |s| < n^{2/3}$ and hence $\deg F \ll n^{2/3}$. Since $\phi(n) \gg n^{1-\epsilon}$ for all $\epsilon > 0$, the assumption that there are infinitely many torsion points (that is we can take n arbitrarily large) implies the identity $f^r = g^s$ as required.

3. Polynomials Vanishing at Roots of Unity

In this section we present a proof of the following strong version of Lang's conjecture for torsion points on a variety V in $\mathbb{C}^m$. We denote by $\mathbb{U}_{\text{tors}}$ be the set of roots of unity.

COROLLARY 3.1. *Let V be an algebraic variety embedded in $\mathbb{C}^m$. There exists an explicitly computable, finite list $\mathcal{B}$ of ℓ_B-by-m integer matrices B, with each $\ell_B \ge 1$, such that if $\zeta \in V(\mathbb{U}_{\text{tors}})$ then $\zeta \in \bigcup_{B \in \mathcal{B}} W_B(\mathbb{U}_{\text{tors}})$ where $W_B = \bigcap_{j=1}^{\ell_B} \{\zeta : \zeta_1^{b_{j,1}} \zeta_2^{b_{j,2}} \cdots \zeta_m^{b_{j,m}} = 1\}$.*

It is not difficult to give an explicit description of $W(\mathbb{U}_{\text{tors}})$—see at the end.

To prove this result we shall develop a simple understanding of vanishing sums of roots of unity– see (Conway and Jones, 1976) and (Lenstra, 1979) for far more. We begin by considering a linear form $a_1X_1 + a_2X_2 + \cdots + a_kX_k$ where each a_i is an integer. We are interested in finding all sets $(\xi_1, \xi_2, \ldots, \xi_k) \in \mathbb{U}_{\text{tors}}^k$ such that $a_1\xi_1 + a_2\xi_2 + \ldots + a_k\xi_k = 0$. We call such a sum *minimal* if no proper vanishing sums of roots of unity subsum equals zero (that is, there does not exist a proper subset I of $\{1, \dots, k\}$ for which $\sum_{i \in I} a_i \xi_i = 0$); it occurs no loss of generality in our calculations to partition any such sum into minimal subsums. Given any such minimal solution there are *equivalent* solutions $(\xi\xi_1, \xi\xi_2, \dots, \xi\xi_k)$ for any root of unity ξ. Two solutions are *equivalent* if they can be partitioned (in the same way) into minimal subsums, where the corresponding subsums are equivalent.

For any set $(\xi_1, \xi_2, \dots, \xi_k) \in \mathbb{U}_{\text{tors}}^k$ there is a minimal $n = n(\xi_1, \xi_2, \dots, \xi_k)$ for which $(\xi_i/\xi_j)^n = 1$ for each pair $1 \leq i, j \leq k$. Note that any minimal sum $\sum_{i=1}^k a_i \xi_i = 0$ is thus equivalent to a minimal solution $\sum_{i=1}^k a_i \xi_i' = 0$ where each $(\xi')^n = 1$, with $n = n(\xi_1, \xi_2, \dots, \xi_k)$. Our key result is the following:

PROPOSITION 3.2. *Suppose that $a_1\xi_1 + a_2\xi_2 + \ldots + a_k\xi_k = 0$ is minimal. Then $n(\xi_1, \xi_2, \dots, \xi_k)$ is squarefree, and if prime p divides n then $p \leq k$. Therefore n divides $N_k := \prod_{p \leq k} p$.*

Given non-zero integers $a_1, a_2, \dots, a_k$, let $\mathbb{X} = \mathbb{X}(a_1, \dots, a_k)$ be the set

$$\{(\xi_1, \dots, \xi_k) : \xi_j^{N_k} = 1 \text{ for each } j, \text{ and } a_1\xi_1 + \ldots + a_k\xi_k = 0\},$$

which is finite and computable, simply by trying all possible values for each ξ_j. One consequence of Proposition 3.2 is the following result:

COROLLARY 3.3. *Suppose $a_1, \dots, a_k \in \mathbb{Z}^*$. For given $(\xi_1, \xi_2, \dots, \xi_k) \in \mathbb{U}_{\text{tors}}^k$ we have $a_1\xi_1 + a_2\xi_2 + \ldots + a_k\xi_k = 0$ if and only if $(\xi_1, \xi_2, \dots, \xi_k)$ is equivalent to an element of $\mathbb{X}$.*

Proof. Given $a_1\xi_1 + a_2\xi_2 + \ldots + a_k\xi_k = 0$, split the sum up into minimal subsums, each one of which (according to the remarks above) is equivalent to one where each ξ_i is an nth root of unity. Moreover n divides $N_\ell = \prod_{p \leq \ell} p$ by Proposition 3.2, where ℓ is the length of the subsum, and the result follows since $\ell \leq k$. On the other hand if $(\xi_1, \xi_2, \dots, \xi_k)$ is equivalent to an element of $\mathbb{X}$ then $a_1\xi_1 + a_2\xi_2 + \cdots + a_k\xi_k = 0$ by the definition of $\mathbb{X}$.

With that preparation we can prove Corollary 3.1:

Proof of Corollary 3.1. An algebraic variety can be described as the set of points in $\mathbb{C}^m$ satisfying certain equations with algebraic coefficients; and this is a subset of the algebraic variety given by the set of points in $\mathbb{C}^m$ satisfying the norms of these equations, which are equations with integer coefficients. So without loss of generality we will assume the coefficients of the polynomials defining V are integers.

Now suppose that

$$f_j(x_1, \dots, x_m) = \sum_{i=1}^{k_j} a_{j,i} x_1^{s_{j,i,1}} x_2^{s_{j,i,2}} \cdots x_m^{s_{j,i,m}} \in \mathbb{Z}[x_1, \dots, x_m]$$

for $1 \le j \le J$. We are interested in $\zeta \in \mathbb{U}_{\text{tors}}^m$ for which $f_j(\zeta) = 0$ for each j; evidently these induce solutions to

$$a_{j,1}\xi_{j,1} + a_{j,2}\xi_{j,2} + \cdots + a_{j,k_j}\xi_{j,k_j} = 0$$

with each $\xi_{j,i} = \zeta_1^{s_{j,i,1}}\zeta_2^{s_{j,i,2}} \cdots \zeta_m^{s_{j,i,m}}$. Now each of these vanishing sums can be partitioned into minimal vanishing subsums; let us relabel one of these minimal vanishing subsums to be $a_1\xi_1 + a_2\xi_2 + \ldots + a_k\xi_k = 0$. As we saw in Proposition 3.2, each $\xi_r/\xi_1 = \zeta_1^{s_{r,1}-s_{1,1}}\zeta_2^{s_{r,2}-s_{1,2}} \cdots \zeta_m^{s_{r,m}-s_{1,m}}$ must be an N_kth root unity, so $\zeta_1^{b_{r,1}}\zeta_2^{b_{r,2}} \cdots \zeta_m^{b_{r,m}} = 1$ where $b_{r,j} = N(s_{r,j} - s_{1,j})$ for each j. We get sets of such vectors b_r for each minimal vanishing subsum (and from each f_j) and we can concatenate these all together to form one large matrix B (with, say, ℓ rows), and so $\zeta \in W_B(\mathbb{U}_{\text{tors}})$.

Finally, since there are only finitely many possible partitions into minimal subsums, the set $\mathcal{B}$ of such matrices B, is finite and computable.

Proof of Proposition 3.2. Write each $\xi_j = e(k_j/n)$ with $0 \le k_j \le n-1$.

Suppose that integer r divides n, and let $\beta_j \equiv k_j \pmod{n/r}$ with $0 \le \beta_j \le n/r - 1$, and $\gamma_j = (k_j - \beta_j)/(n/r)$ so that $0 \le \gamma_j \le r-1$. Thus $\xi_j = e(\beta_j/n)e(\gamma_j/r)$. Now, for each $0 \le i \le r-1$ and $0 \le \ell \le n/r - 1$, let $A_{i,\ell}$ be the sum of the a_j with $\beta_j = \ell$ and $\gamma_j = i$ so that

$$0 = a_1\xi_1 + a_2\xi_2 + \ldots + a_k\xi_k = \sum_{j=0}^{k} a_j e(\beta_j/n)e(\gamma_j/r) = \sum_{\ell=0}^{n/r-1}\left(\sum_{i=0}^{r-1} A_{i,\ell}e(i/r)\right)e(\ell/n).$$

Let $r = r(n) = \prod_{p|n} p$ and recall that $[\mathbb{Q}(e(1/n)) : \mathbb{Q}(e(1/r))] = n/r$ (by elementary Galois theory) and so $e(\ell/n)$, $0 \le \ell \le n/r - 1$ are linearly independent over $\mathbb{Q}(e(1/r))$. In particular this implies that each of the subsums $\sum_{i=0}^{r-1} A_{i,\ell}\, e(i/r) = 0$ above, which contradicts our assumption of minimality, unless $A_{i,\ell} = 0$ for all i for all $\ell \ne \ell_0$ for some ℓ_0; in other words $\beta_j = \ell_0$ for all j. But then $\xi_i/\xi_j = e(\ell_0/n)e(\gamma_j/r)/e(\ell_0/n)e(\gamma_j/r) = e((\gamma_i - \gamma_j)/r)$ and so $n(\xi_1, \xi_2, \ldots, \xi_k)$ divides r. Thus $n = r(n)$ is squarefree.

Since n is squarefree we may write $n = mp$ with $(m, p) = 1$. Then, by the Chinese Remainder theorem there exists $0 \le \beta_j \le p-1$ and $0 \le \gamma_j \le m-1$ such that $k_j \equiv m\beta_j \pmod{p}$ and $k_j \equiv p\gamma_j \pmod{m}$ and thus $\xi_j = e(\beta_j/p)e(\gamma_j/m)$. Letting $A_{i,\ell}$ now be the sum of the a_j with $\beta_j = \ell$ and $\gamma_j = i$ we obtain

$$0 = a_1\xi_1 + a_2\xi_2 + \cdots + a_k\xi_k = \sum_{j=0}^{k} a_j e(\beta_j/p)e(\gamma_j/m) = \sum_{\ell=0}^{p-1}\left(\sum_{i=0}^{m-1} A_{i,\ell}e(i/m)\right)e(\ell/p).$$

Recall that $[\mathbb{Q}(e(1/n)) : \mathbb{Q}(e(1/m))] = p-1$ (by elementary Galois theory), so that the only linear dependencies between $e(\ell/p)$, $0 \le \ell \le p-1$, over

$\mathbb{Q}(e(1/m))$, are multiples of $\sum_{\ell=0}^{p-1} e(\ell/p) = 0$. Therefore from the equation above we see that $\sum_{i=0}^{m-1} A_{i,\ell} e(i/m) = \lambda$ for some $\lambda \in \mathbb{Q}(e(1/m))$. Evidently $\lambda \neq 0$ else, by the argument from the paragraph above we see that $n \mid m$. Therefore for each ℓ there exists i with $A_{i,\ell} \neq 0$ and in particular some $j = j_\ell$ with $\beta_{j_\ell} = \ell$; and so $p \leq k$ as claimed.

3.1. DETERMINING $W_B(\mathbb{U}_{\mathrm{tors}})$

Suppose that the ℓ-by-m integer matrix B is given and we write each $\zeta_j = e(v_j)$, so the points in W_B correspond exactly to those $v \in (\mathbb{Q}/\mathbb{Z})^m$ satisfying $Bv \equiv 0 \pmod 1$. Note that if $y \in B^\perp(\mathbb{Q}) \pmod 1$ then $By \equiv 0 \pmod 1$, so we call two solutions v, v' *equivalent* if $v - v' \in B^\perp(\mathbb{Q}) \pmod 1$. We will prove that there are no more than finitely many inequivalent solutions, which are effectively computable:

We wish to use the tools of linear algebra to solve this equation but there are many zero divisors in $\mathbb{Q} \pmod 1$ (indeed if $a/q \in \mathbb{Q}$ then $q \cdot (a/q) \equiv 0 \pmod 1$), so we avoid any division! In Gaussian elimination one diagonalizes as much of the matrix as possible, dividing non-zero elements in a given row by the "pivot element" (that is if $B_{1,1} \neq 0$ is the pivot element then one replaces the current row i by the current row i minus $B_{i,1}/B_{1,1}$ times the first row). This can be reworked to avoid division simply by introducing multiples (that is we replace the current row i by $B_{1,1}$ times the current row i minus $B_{i,1}$ times the first row). Note that any solution of the original linear algebra problem is also a solution of the new problem; and vice-versa whenever $B_{1,1}$ is invertible, though if this is not so (as may be the case here) this process may well introduce several bogus solutions. Nonetheless at the end of the Gaussian elimination process we have an l-by-m integer matrix B' (with $l \leq \ell$ after deleting rows of 0s), in which the left-most l-by-l submatrix is diagonal with non-zero diagonal entries (if necessary by swapping various rows and columns), for which $B'v \equiv 0 \pmod 1$. Solving this is easy: there are $m - l$ free variables $v_{l+1}, v_{l+2}, \ldots, v_m$ and, writing $\beta_i = B'_{i,i}$, we have $v_i \equiv (u_i - \sum_{j=l+1}^{m} B'_{i,j} v_j)/\beta_i \pmod 1$, where u_i is any integer with $0 \leq u_i \leq \beta_i - 1$.

For $l + 1 \leq j \leq m$ let y_j be the vector with ith entry $-B'_{i,j}/\beta_i$ for $1 \leq i \leq \ell$, and $\delta_{i,j}$ otherwise (where δ is the Dirac delta function). The solutions to $B'v \equiv 0 \pmod 1$ all take the form $v = u + \sum_{j=l+1}^{m} v_j y_j$ where $u \in U'$ a finite computable set. If we trace through the proof above then we find that $By_j = 0$ for each j, that is each $y_j \in B^\perp$. Thus there is a set U of representatives of the equivalence classes of solutions inside U' which can be determined by testing whether they satisfy $Bu \equiv 0 \pmod 1$.

References

Conway, J. H. and Jones, A. J. (1976) Trigonometric diophantine equations on vanishing sums of roots of unity, *Acta Arith.* **30**, 229–240.

Lang, S. (1965) Division points on curves, *Ann. Mat. Pura Appl. (4)* **70**, 229–234.

Lenstra, Jr., H. W. (1979) Vanishing sums of roots of unity, In *Proc. Bicentennial Congress Wiskundig Genootschap*, Vol. 101 of *Math. Centre Tracts*, Vrije Univ. Amsterdam, 1978, pp. 249–268, Math. Centrum, Amsterdam.

Strombergsson, A. and Venkatesh, A. (2005) Small solutions to linear congruences and Hecke equidistribution, *Acta Arith.* **118**, 41–78.

Tzermias, P. (2000) The Manin–Mumford conjecture: a brief survey, *Bull. London Math. Soc.* **32**, 641–652.

Ullmo, E. (2006) Manin–Mumford, André–Oort, the equidistribution point of view, in this book.

Venkatesh, A. (2006) Spectral theory of automorphic forms: a very brief introduction, in this book.

THE DISTRIBUTION OF ROOTS OF A POLYNOMIAL

Andrew Granville
Université de Montréal

1. Introduction

How are the roots of a polynomial distributed (in $\mathbb{C}$)? The question is too vague for if one chooses one's favourite complex numbers $z_1, z_2, \ldots, z_d$ then the polynomial $\prod_{j=1}^{d}(x - z_j)$ has its roots at these points. However if one looks at polynomials that arise frequently then one finds that certain patterns emerge. Take for example $x^n - 1$. Here the roots are equidistributed around the unit circle, at the points $\{e(j/n) : 0 \le j \le n-1\}$, and the larger n, the more points one has, and the denser they become. (Throughout this article, $e(t) := e^{2i\pi t}$.)

In terms of measure, write $\mu_{\{f\}} = (1/n)\sum_{j=1}^{n}\delta_{z_j}$ for a polynomial with (not necessarily distinct) roots $z_1, z_2, \ldots, z_d$, where δ is the Dirac delta-measure. Let $\nu_{\{|z|=1\}}$ be the Haar measure on the unit circle (that is, uniform distribution). Then we have $\lim_{n\to\infty}\mu_{\{x^n-1\}} = \nu_{\{|z|=1\}}$ (that is, convergence in the sense of "weak convergence").

Another interesting example is $(x-1)^n$; in this case all the roots are at the same point on the unit circle, 1; and so $\lim_{n\to\infty}\mu_{\{(x-1)^n\}} = \delta_1$. One more example is $x^n - 2$. Here the roots are again equidistributed in angle where, as n gets larger, the more points one has, and the more uniformly distributed they become. But there is more than that. As $n \to \infty$ we have $2^{1/n} \to 1$, so all of the roots get closer and closer to the unit circle as $n \to \infty$. Therefore $\lim_{n\to\infty}\mu_{\{x^n-2\}} = \nu_{\{|z|=1\}}$.

So what distinguishes those sequences of polynomials for which the limiting measure of the roots is the Haar measure on the unit circle? The most obvious difference if one compares polynomials $x^n - a$ where $a^{1/n} \to 1$ as $n \to \infty$, and polynomials like $(x-1)^n$ is that the latter has coefficients whose size grow exponentially in n, whereas the former do not. The main point of this section is to prove a result along these lines: "If the coefficients of $f(x) \in \mathbb{C}[x]$ are not too large then $\mu_{\{f\}}$ is not far from $\nu_{\{|z|=1\}}$". Obviously we need to be more precise than this, but we run into a tricky question: What is the best measure of the size of the coefficients of a polynomial? There are

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 93–102.

several options used in the literature, and it is known that they do not differ in size by much – however the "by much" can be as large as exponential in the degree of f which is too much for our application. The first result in this direction, due to Erdős and Turán (Erdős and Turán, 1950) used the renormalized 1-norm. If $f(x) = a_d \prod_{j=1}^d (x - \alpha_j) = \sum_{j=0}^d a_j x^j$ where $a_d a_0 \neq 0$ then

$$L(f) := \frac{1}{(|a_d|\,|a_0|)^{1/2}} \left(\sum_{j=0}^{d} |a_j| \right).$$

Progress in arithmetic has suggested that the most natural height is the Mahler measure

$$M(f) := |a_d| \prod_{j=1}^{d} \max\{1, |\alpha_j|\};$$

this has several advantages, one of which is that if for a given algebraic number α we take f to be the minimum polynomial for α (over $\mathbb{Q}$) and let $M(\alpha) = M(f)$, then the renormalized height $h(\alpha) := (1/d)\log M(\alpha)$ is simple to use in calculations without reference to the smallest field to which α belongs. Note that if $f^*(x) = x^d f(1/x)$ then $L(f^*) = L(f)$ and $M(f^*) = M(f)$.

Jensen's formula gives an analytic interpretation of Mahler's measure:

$$M(f) = \exp\left(\int_0^1 \log\left|f(e(t))\right| dt \right);$$

and so $M(f) \leq \max_t \left|f(e(t))\right|$. Now, note that $\left|f(e(t))\right| \leq \sum_{j=0}^d |a_j|\,|e(jt)| \leq \sum_{j=0}^d |a_j| = L(f)(|a_d|\,|a_0|)^{1/2}$ and so $M(f) \leq L(f)(|a_d|\,|a_0|)^{1/2}$, which yields

$$\begin{aligned}\prod_{j=1}^{d} \max\{|\alpha_j|, |1/\alpha_j|\} &= \prod_{j=1}^{d} \max\{1, |\alpha_j|\} \prod_{j=1}^{d} \max\{1, 1/|\alpha_j|\} \\ &= \frac{M(f)}{|a_d|} \cdot \frac{M(f^*)}{|a_0|} \leq \frac{(L(f)(|a_d|\,|a_0|)^{1/2})^2}{|a_d|\,|a_0|} = L(f)^2.\end{aligned}$$

We deduce the following result:

LEMMA 1.1. *Suppose that $f_1, f_2, \ldots$ is a sequence of polynomials, where f_d has (not necessarily distinct) roots $\alpha_{d,1}, \alpha_{d,2}, \ldots \alpha_{d,d}$, all non-zero. If $L(f_d) = e^{o(d)}$ as $d \to \infty$ then $|\alpha_{d,j}| = 1 + o(1)$ for $\{1 + o(1)\}d$ values j, $1 \leq j \leq d$.*

This shows that most of the roots come in towards the unit circle. Now we wish to show that they are uniformly distributed around the circle. For a given polynomial f write the roots as $\alpha_j = r_j e(\varphi_j)$ with each $r_j \in \mathbb{R}^+$. For $0 \leq \alpha < \beta \leq 1$ define

$$N_f(\alpha, \beta) = \#\{j : 1 \leq j \leq d \text{ such that } \alpha \leq \{\varphi_j\} < \beta\}.$$

PROPOSITION 1.2 (Erdős and Turán, 1950). *For any polynomial f of degree $d > 1$, and any $0 \le \alpha < \beta \le 1$, we have*

$$|N_f(\alpha,\beta) - (\beta - \alpha)d| \le 8\sqrt{d \log L(f)}.$$

Combining these two results we immediately deduce the result we had guessed at earlier:

THEOREM 1.3. *Suppose that $f_1, f_2, \ldots$ is a sequence of polynomials in $\mathbb{C}[x]$ where f_d has degree d and $f_d(0) \ne 0$. If $L(f_d) = e^{o(d)}$ as $d \to \infty$ then*

$$\lim_{d\to\infty} \mu_{\{f_d(x)\}} = \nu_{\{|z|=1\}}.$$

(We stress that this limit is in the sense of "weak convergence" of measures.)

Erdős and Turán's proof of Proposition 1.2 boils down to the following optimization result.

LEMMA 1.4. *Fix $\gamma \in [0,1)$. Suppose that $g(x)$ has degree d with all of its roots on the unit circle, and that $N_g(0,\gamma) = [\gamma d]+2\Delta+1$. Then* $\max_t \bigl|g(e(t))\bigr| \ge \exp(\Delta^2/4(d+1))$.

Deduction of Proposition 1.2 *from Lemma* 1.4. Given $f(x) = a_d \prod_{j=1}^d (x - \alpha_j) = \sum_{j=0}^d a_j x^j$ take $g(x) = \prod_{j=1}^d (x - e(\varphi_j))$ so that $N_f(0,\gamma) = N_g(0,\gamma)$. Consider the inequality

$$\begin{aligned}\frac{|re(\varphi) - e(t)|^2}{r} &= r + \frac{1}{r} - 2\cos(2\pi(\varphi - t)) \\ &\ge 2 - 2\cos(2\pi(\varphi - t)) = |e(\varphi) - e(t)|^2.\end{aligned}$$

Multiply this over the roots $r_j e(\varphi_j)$ of f, to obtain

$$\bigl|g(e(t))\bigr|^2 \le \frac{\bigl|f(e(t))\bigr|^2}{a_d^2 \prod_j |r_j|} \le \frac{L(f)^2 |a_0 a_d|}{|a_0 a_d|} = L(f)^2$$

since $\bigl|f(e(t))\bigr| \le L(f)|a_0 a_d|^{1/2}$ as we established above, and so $\bigl|g(e(t))\bigr| \le L(f)$. Combining this with Lemma 1.4, we deduce that

$$N_f(0,\gamma) \le [\gamma d] + 1 + 4\sqrt{(d+1)\log L(f)} \le \gamma d + 8\sqrt{d \log L(f)}$$

since $L(f) \ge (|a_0| + |a_d|)/|a_0 a_d|^{1/2} \ge 2$.

Let $h(x) = \sum_{j=0}^d \overline{a_j} x^j$ so that $L(h) = L(f)$ and $N_f(\gamma,1) = N_h(0,1-\gamma)$. Therefore, by the above,

$$N_f(0,\gamma) = d - N_f(\gamma,1) = d - N_h(0,1-\gamma) \ge \gamma d - 8\sqrt{d \log L(h)} = \gamma d - 8\sqrt{d \log L(f)}.$$

Now let $h(x) = f(e(\alpha)x) = \sum_j b_j x^j$ so that $b_j = e(j\alpha)a_j$ for each j, with $L(h) = L(f)$, and $N_h(0,\gamma) = N_f(\alpha,\beta)$ where $\gamma = \beta - \alpha$. Thus we may assume, without loss of generality, that $\alpha = 0$ by replacing f with h. The result then follows from the previous two displayed equations.

Their proof of Lemma 1.4 involves several ingenious arguments, blending facts then well-known about polynomials. In the next section we will see a different and complete proof of a related result so here we will just give a

Sketch of their proof of Lemma 1.4. The idea is to understand the optimal polynomial; that is, g satisfying the hypothesis for which $\max_t \left|g(e(t))\right|$ is minimal. So the first thing they do is to show that it takes its maximal value at some point in-between each pair of roots of $g(x)$ inside the arc in question. Next they apply a result of Turán which says that there cannot be a zero of $g(x)$ at a distance less than $\pi/2d$ from one of these maximal points; which implies that at least 2Δ roots must lie at the endpoints of the interval, and so at least one endpoint has a zero with multiplicity $\geq \Delta$. We know from basic complex analysis that a polynomial with a root of high multiplicity must get large, and thus they obtain their lower bound.

2. Algebraic Numbers

Theorem 1.3 is a purely analytic result, in that there are no algebraic requirements on f. It is of more interest in arithmetic to have such requirements; for example if we insist that all of the coefficients of f are integers then the roots of f are the union of various complete sets of conjugates of certain algebraic numbers. In this circumstance Bilu proved an arguably stronger result than Theorem 1.3 (in that it involves $M(f)$ rather than $L(f)$) with a better motivated proof. Also, as we shall see in the next section, it generalizes to higher dimension in a beautiful way.

For a compactly supported measure μ on $\mathbb{C}$ we define the *energy* by

$$E(\mu) := -\int\int \log|z-w|\, d\mu_z\, d\mu_w.$$

If μ is finitely supported, at $\{\alpha_1, \dots, \alpha_d\}$, then define

$$E'(\mu) := -\sum_{i\neq j} \mu(\alpha_i)\mu(\alpha_j)\log|\alpha_i - \alpha_j| \quad \text{and} \quad \|\mu\| = \left(\sum_i \mu(\alpha_i)^2\right)^{1/2}.$$

Note that $E'(\mu)$ is not the same as $E(\mu)$ since we miss out the $i = j$ terms (and note that $E(\mu) = \infty$ by including them). We quote a couple of useful results on measures from the literature:

LEMMA A. *If $\{\mu_d\}_{d=1,2,...}$ have finite support with $||\mu_d|| \to 0$ as $d \to \infty$ and where the μ_d weakly converge to μ, then*

$$E(\mu) \leq \liminf E'(\mu_d).$$

LEMMA B. *If $K \subset \mathbb{C}$ is compact then there exists a unique measure $\nu = \nu_K$ for which $E(\nu)$ is minimized over all measures ν whose support is a subset of K (we call ν_K the* equilibrium measure *of K). If $K = \{|z| = 1\}$ is the unit circle then $E(\nu_K) = 0$.*

Now suppose $f(x) \in \mathbb{Z}[x]$ has distinct roots $\alpha_1, \ldots, \alpha_d$ and lead coefficient a_d. The discriminant of f is a non-zero integer, $\mathrm{Disc}(f) := a_d^{2d-2} \prod_{i \neq j}(\alpha_i - \alpha_j)$. Therefore

$$\begin{aligned} 0 \leq \frac{1}{d^2} \log(\mathrm{Disc}(f)) &= \frac{2d-2}{d^2} \log|a_d| + \sum_{i \neq j} \frac{1}{d^2} \log|\alpha_i - \alpha_j| \\ &\leq \frac{2}{d} \log M(f) - E'(\mu_{\{f\}}), \end{aligned}$$

and so $E'(\mu_{\{f\}}) \leq (2/d) \log M(f)$. We deduce

THEOREM 2.1. *Suppose that $f_1, f_2, \ldots$ is a sequence of polynomials in $\mathbb{Z}[x]$ where f_d has degree d and $f_d(0) \neq 0$. If $M(f_d) = e^{o(d)}$ as $d \to \infty$ then*

$$\lim_{d \to \infty} \mu_{\{f_d(x)\}} = \nu_{\{|z|=1\}}.$$

Proof. As $f_d(x) \in \mathbb{Z}[x]$ we see $\prod_j \max\{1, |\alpha_j|\} \leq M(f_d) = e^{o(d)}$ and $\prod_j \max\{1, 1/|\alpha_j|\} \leq M(f_d^*) = M(f_d) = e^{o(d)}$, so $\mu_{\{f_d\}}$ is converging to some measure on the unit circle. Now since this is compact there must be some subsequence of the f_d such that $\mu_{\{f_d\}}$ converges weakly to some limit, call it μ, supported on $\{|z| = 1\}$, on that subsequence. But since $||\mu_d|| = 1/\sqrt{d} \to 0$ as $d \to \infty$ we may apply Lemma A to deduce that $E(\mu) \leq \liminf E'(\mu_d) \leq \liminf(2/d) \log M(f_d) = 0$. On the other hand $E(\mu) \geq E(\nu_{\{|z|=1\}}) = 0$ by Lemma B, and so $E(\mu) = 0$. However this implies that $\mu = \nu_{\{|z|=1\}}$ by Lemma B, and this is true for any convergent subsequence. From the above compactness argument it is clear that all f_d belong to some convergent subsequence, and since they all have this same limiting measure the result follows.

If α is an algebraic number with minimum polynomial f we define $M(\alpha) = M(f)$; Theorem 2.1 can easily be reformulated in terms of a sequence of algebraic numbers α_d.

3. In *k* Dimensions: the Bilu Equidistribution Theorem

Bilu (Bilu, 1997) went on from here to consider whether the conjugates of different algebraic numbers of small height are distributed independently. In other words if we are given sequences $\{(\alpha_i,\beta_i) : i = 1,2,\ldots\}$ of algebraic numbers, where both α_i and β_i have degree i over the rationals and both have small height (as above) then what is the joint distribution of the conjugates? In other words, let H_i be the subgroup of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ which gives all distinct pairs of conjugates $S_i := \{(\alpha_i^\sigma,\beta_i^\sigma) : \sigma \in H_i\}$, and then consider the measure

$$\mu_{S_i} = \frac{1}{|S_i|}\sum_\sigma \delta_{\{\alpha_i^\sigma,\beta_i^\sigma\}}.$$

Evidently $S_i \subset \overline{\mathbb{Q}}^2$ and, in the limit we know (from the previous section) that these measures are supported on the two dimensional torus $\mathbb{U}^2$ where $\mathbb{U} := \{z \in \overline{\mathbb{Q}}^* : |z| = 1\}$. Then the question is whether they have a limit and, if so, whether that limit is the Haar measure. In other words, writing $\alpha_i^\sigma = c_i e(\varphi_{i,\sigma})$ and $\beta_i^\sigma = b_i e(\theta_{i,\sigma})$, we ask whether for any $0 \le u_j < v_j \le 1$,

$$\#\{\sigma \in H_i : u_1 \le \varphi_{i,\sigma} < v_1 \quad \text{and} \quad u_2 \le \theta_{i,\sigma} < v_2\} \sim (v_1 - u_1)(v_2 - u_2)|H_i|$$

as $i \to \infty$? There are some obvious cases in which this cannot be true: for example if $\alpha_i = \beta_i$ for each i, or $\alpha_i\beta_i = 1$ for each i. So the conjecture becomes that the pairs should be uniformly distributed on $\mathbb{U}^2$, unless they belong to some obvious family of exceptions and it is now necessary to try to determine the correct formulation of the exceptional set:

The k-dimensional algebraic torus $\mathbb{T}_k$ is isomorphic to $(\overline{\mathbb{Q}}^*)^k$, and a *torsion subvariety* of $\mathbb{T}_k$ is a translate of a subtorus by a torsion point. We will call a sequence of points $\alpha_1, \alpha_2, \ldots \in \mathbb{T}_k$ *strict* if there are only finitely many such points in any proper torsion subvariety. Bilu's result works for strict sequences of small height, though first we need to define height here:

Given $\gamma \in \overline{\mathbb{Q}}^*$ define $\deg(\gamma)$ to be the degree of the minimum polynomial of γ, and let $H(\gamma) = M(\gamma)^{1/\deg(\gamma)}$ (thus the condition in Theorem 2 can be rewritten as $H(\gamma_d) = 1 + o(1)$ if f_d is the minimum polynomial of γ_d). For $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_k) \in \mathbb{T}_k$, define $\deg(\gamma) = \min_i \deg(\gamma_i)$, and $H(\gamma) = \prod_i H(\gamma_i)$.

THEOREM 3.1 (Bilu, 1997). *Suppose that $\alpha_1, \alpha_2, \ldots \in \mathbb{T}_k$ is a strict sequence with* $\deg(\alpha_d) \ge d$ *for each d, and $H(\gamma_d) = 1 + o(1)$ as $d \to \infty$ (we call this last condition* small height). *Then*

$$\lim_{d\to\infty} \mu_{\{\alpha_d\}} = \nu_{\mathbb{U}^k}.$$

Sketch of Proof. As we noted above, almost all conjugates of α_d are getting closer and closer to $\mathbb{U}^k$ as $d \to \infty$ and so, by compactness, there must be some subsequence that tends to a limiting measure (call it ν).

For any non-trivial character $\chi : \mathbb{T}_k \to \overline{\mathbb{Q}}$ the sequence $\chi(\alpha_d)$ is strict and has small height. Thus applying Theorem 2.1 to our subsequence we see that $\chi_*(\nu) = \nu_{\mathbb{U}}$ (where χ_* should be interpreted as the action of χ on the support of the measure, and thus the measure). But this is true for any non-trivial character χ and so ν must be the Haar measure on $\mathbb{U}_k$, namely $\nu_{\mathbb{U}^k}$. But this is true for any convergent subsequence and so for the whole sequence (by the same argument as in the proof of Theorem 2.1).

This beautiful result has many powerful consequences. Most famous, perhaps, is

COROLLARY 3.2 ((Zhang, 1995)). *Suppose that $X \subset \mathbb{T}_k$ is Zariski closed. Let W be the union of the torsion subvarieties that lie entirely in X. There exists a constant $c(X) > 1$ such that if $\alpha \in X \setminus W$ then $H(\alpha) > c(X)$.*

Proof. Northcott's theorem tells us that there are only finitely many algebraic numbers of given degree below a certain height. So if Zhang's theorem is false then there is an strict sequence $\alpha_1, \alpha_2, \ldots \in \mathbb{T}_k$ with $\deg(\alpha_d) \geq d$ for each d, and $H(\alpha_d) = 1 + o(1)$ as $d \to \infty$. By Bilu's Theorem (Theorem 3.1) these become equidistributed around $\mathbb{U}^k$, and so, as $X(\mathbb{T}_k)$ is closed (since X is Zariski closed), thus $\mathbb{U}^k \subset X(\mathbb{T}_k)$. However $\mathbb{U}^k$ is Zariski dense (as may be proved by induction on k) and so $X(\mathbb{T}_k) = \mathbb{T}_k$ in which case $W = X$ and the result is trivial.

This result was proved by Szpiro, Ullmo and Zhang for points on an abelian variety by rather different means, something that will be discussed by Ullmo (Ullmo, 2006) in a subsequent section.

It is perhaps a little difficult to understand Zhang's theorem, so let's examine a special case, the solutions to $x + y = 1$ in algebraic numbers. There are four torsion solutions $1+0 = 0+1 = 1$ and $e(1/6)+e(5/6) = e(5/6)+e(1/6) = 1$, so we now investigate solutions omitting these: So suppose that we have a solution $\alpha + (1 - \alpha) = 1$ with $H(\alpha)H(1 - \alpha)$ small. Note that if $H(\beta)$ is small then most conjugates β^σ of β must be close to the unit circle, that is $|\beta^\sigma| \approx 1$. Thus for most conjugates α^σ of α we have $|\alpha^\sigma|, |1 - \alpha^\sigma| \approx 1$. The circles of radius 1 centered at 0 and 1 only intersect at $e(1/6)$ and $e(5/6)$, so we must have $\alpha^\sigma \approx e(\pm 1/6)$ whence $1 - \alpha^\sigma \approx e(\mp 1/6)$ for almost all σ. In such cases $(\alpha^\sigma)^2 - \alpha^\sigma + 1 = 1 - \alpha^\sigma(1 - \alpha^\sigma) \approx 0$, and thus the norm of $\alpha^2 - \alpha + 1$ over $\mathbb{Q}$ is very small, whereas it should be a non-zero integer (and thus ≥ 1 in absolute value). Formalizing and refining this argument, Zagier (Zagier, 1993) was able to show that in any non-torsion algebraic solution of

$x+y=1$ we have $H(x)H(y) \geq ((1+\sqrt{5}/2)^{1/2}$. It is amusing to try to develop an analogous argument for other varieties (see for example (Bombieri and Zannier, 1995)).

4. Lower Bounds on Heights

Kronecker established a result on roots of unity on the unit circle which can be re-interpreted as stating that for any integer $d \geq 1$, if $M(\alpha) > 1$ then there exists a constant $\delta(d) > 0$ such that $M(\alpha) \geq 1 + \delta(d)$ for all α of degree d. In 1933 Lehmer (Lehmer, 1933) made the extraordinary conjecture that $\delta(d) \geq \delta(10) = .1762808\ldots$ obtained from the example where α is a root of $x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1$.

In 1979 Dobrowolski (Dobrowolski, 1979) showed that one can take $\delta(d) = (1-\epsilon)(\log\log d/\log d)^3$ but this has not been much improved subsequently.

5. Compact Sets with Minimal Energy

Let K be a compact subset of the complex plane and suppose that $E(\nu_K) = 0$. One example is where K is the unit circle but there are many other interesting examples besides (for example the line segment $[-2,2]$). Rumely (Rumely, 1999) showed that Bilu's result, our Theorem 2.1, can be extended when appropriately reformulated to any such K (note though that Rumely prefers to work with the *capacity* of K, which is given by $\exp(-E(\nu_K))$, and is thus 1 in this case).

I cannot resist at least mentioning the first few results of capacity theory as they motivate some of the ideas in this article, which link several obvious notions: First we have that if $\alpha_1,\ldots,\alpha_d \in K$ is the support of measure $\mu = \mu_\alpha$ where $\mu_\alpha(\alpha_j) = 1/d$ then

$$E(\nu_K) = \lim_{d\to\infty} \min_{\alpha_1,\ldots,\alpha_d\in K} E'(\mu_\alpha).$$

The right-side here is the logarithm of what Fekete called the *transfinite diameter* of a compact set K. From this one can deduce that ν_K is supported only on the outer boundary of K.

Second we have

$$E(\nu_K) = \lim_{d\to\infty} \min_{\substack{\text{monic } f(z)\in\mathbb{C}[z] \\ \deg f=d}} \frac{1}{d} \sup_{z\in K} \log|f(z)|;$$

the right-side here is the logarithm of the *Chebyshev constant* of K.

The Mahler measure used above, when K is the unit disk, can be reinterpreted as a product of parts. The infinite part is the product of terms $\max\{1, |\alpha|\}$, the "local parts" the prime powers dividing $|a_d|$. In defining an appropriate height function for more general K we give the same definition for the local parts, but the infinite part is now defined as the exponential of minus the *potential* function for K. More explicitly

$$M_K(\alpha) := |a_d| \prod_{\sigma} G(\alpha^{\sigma}, \infty; K),$$

where our Green's function, $G(z, \infty; K)$, is defined as $E(\nu_K) + \int_K \log(|z - w|) d\nu_K(w)$. (Note that $G(z, \infty; K) = \max(0, \log(|z|)$ when $K = D(0, 1)$.) We interpret $M_K(f) = M_K(\alpha)$ for any root α of irreducible f; and $M_K(fg) = M_K(f)M_K(g)$.

Now, as in the proof of Theorem 2.1, we have, using the above and the well-known fact that $G(z, \infty; K) \geq 0$ for all z,

$$0 = E(\nu_K) \leq \liminf E'(\mu_\alpha) \leq \liminf \frac{2}{d} \log M_K(\alpha) = 0.$$

In this way, Rumely proved:

THEOREM 5.1. *Fix a compact set K with $E(\nu_K) = 0$. Suppose that $f_1, f_2, \ldots$ is a sequence of polynomials in $\mathbb{Z}[x]$ where f_d has degree d and $f_d(0) \neq 0$. If $M_K(f_d) = e^{o(d)}$ as $d \to \infty$ then*

$$\lim_{d\to\infty} \mu_{\{f_d(x)\}} = \nu_K.$$

Acknowledgements

Thanks to Ashkan Nikeghbali for pointing out to me, during the workshop, that a result like Lemma 1.1 is easily deduced from the method in (Hughes and Nikeghbali, 2006). Also to Robert Rumely for his remarks (and corrections) on these notes. Parts of this lecture were derived from notes kindly lent to me by Yuri Bilu.

References

Bilu, Y. (1997) Limit distribution of small points on algebraic tori, *Duke Math. J.* **89**, 465–476.

Bombieri, E. and Zannier, U. (1995) Algebraic points on subvarieties of $\mathbb{G}_m^n$, *Internat. Math. Res. Notices* **1995**, 333–347.

Dobrowolski, E. (1979) On a question of Lehmer and the number of irreducible factors of a polynomial, *Acta Arith* **34**, 391–401.

Erdős, P. and Turán, P. (1950) On the distribution of roots of polynomials, *Ann. Math.* **51**, 105–119.

Hughes, C. and Nikeghbali, A. (2006) The zeros of random polynomials cluster uniformly near the unit circle, to appear.

Lehmer, D. H. (1933) Factorization of certain cyclotomic functions, *Ann. Math.* **34**, 461–479.

Rumely, R. (1999) On Bilu's equidistribution theorem, *Contemp. Math.* **237**, 159–166.

Ullmo, E. (2006) Manin–Mumford, André–Oort, the equidistribution point of view, in this book.

Zagier, D. (1993) Algebraic numbers close to both 0 and 1, *Math. Comp.* **61**, 485–491.

Zhang, S. (1995) Positive line bundles on arithmetic varieties, *J. Amer. Math. Soc.* **8**, 187–221.

MANIN–MUMFORD, ANDRÉ–OORT, THE EQUIDISTRIBUTION POINT OF VIEW

Emmanuel Ullmo
Université Paris-Sud

1. Introduction

These notes were prepared for the 2005 Summer School "*Equidistribution in number theory*" organized by Andrew Granville and Zeev Rudnick in Montréal. It's a pleasure to thank them for the opportunity of giving these lectures. The aim of this text is to describe the conjectures of Manin–Mumford, Bogomolov and André–Oort from the point of view of equidistribution. This includes a discussion of equidistribution of points with small heights of CM points and of Hecke points. We tried also to explain some questions of equidistribution of positive dimensional "special" subvarieties of a given variety.

The assignment by the organizer was to try to present a large overview but to avoid using technical language. For example I was not allowed to use the following notions (I quote the organizers):

1. Adeles (avoid them like the plague!).
2. Shimura varieties.
3. Semisimple groups.
4. Arakelov theory.

I was unable to fill out all the requirements but I tried to focus on significant examples and the presentation of the general picture in a coherent view. Complete proofs are given in only a small amount of simplified cases when it can help the reader improve his intuition on the general case. The main statements are only sketched and the material of these notes covers much more than it's possible to present in a few hours of lectures. We hope that this text will complement the lectures and will be of some help for the reader interested in the understanding of these topics in a deeper way.

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 103–138.

2. Informal Examples of Equi-Distribution

2.1. PRELIMINARY RESULTS FROM MEASURE THEORY

Let X be a metric space and $\mathcal{P}(X)$ the set of Borel probability measures on X. Let $C(X)$ be the set of bounded continuous functions on X. We say that a sequence $\mu_n \in \mathcal{P}(X)$ is weakly convergent to $\mu \in \mathcal{P}(X)$ if for all $f \in C(X)$

$$\mu_n(f) = \int_X f\, d\mu_n \to \mu(f) = \int_X f\, d\mu \quad \text{as } n \to \infty.$$

We'll write $\mu_n \to \mu$ in this case.

We define the weak* topology on $\mathcal{P}(X)$ as the smallest topology making each of the maps $\mu \to \mu(f) = \int f\, d\mu$ $(f \in C(X))$ continuous.

PROPOSITION 2.1. *Suppose that X is a compact metric space. Then μ_n weakly converges to μ if and only if μ_n converges to μ in the weak* topology. The space $\mathcal{P}(X)$ is metrisable and compact for the weak* topology: if $\mu_n \in \mathcal{P}(X)$ is a sequence then there exists a weakly convergent subsequence.*

A useful way of proving some equidistribution properties is given by Weyl's criterion:

PROPOSITION 2.2. *Let X be a compact metric space. Let $\phi_n \in C(X)$ be a sequence with the property that their linear combinations are dense in $C(X)$ (endowed with the usual norm $\|f\| = \sup_{x\in X} |f(x)|$). Then $\mu_n \to \mu$ if and only if for all $m \in \mathbb{N}$, $\mu_n(\phi_m) \to \mu(\phi_m)$.*

If E is a finite subset of X we define $\mu_E \in \mathcal{P}(X)$ as

$$\mu_E = \frac{1}{|E|} \sum_{x\in E} \delta_x \tag{1}$$

where δ_x denotes the Dirac measure supported at x. We say that a sequence E_n of finite subsets of X is equidistributed for $\mu \in \mathcal{P}(X)$ (or μ-equidistributed) if $\mu_{E_n} \to \mu$.

When X is not compact it's sometimes possible to adapt Weyl's criterion. For modular curves or more generally Shimura varieties, L^2-techniques (spectral decomposition) are used to prove the equidistribution of Hecke points or CM points.

EXERCISE 2.3. Let $X = \mathbb{C}^*$ and E_n be the set of nth roots of unity. Then E_n is equidistributed for the normalized measure $d\alpha/2\pi$ supported on the unit circle.

EXERCISE 2.4. Prove the last assertion using Proposition 2.2. Let n be a integer, ζ_n a primitive n-roots of unity. Let E'_n be the set of Galois conjugate of ζ_n. Using the irreducibility of the cyclotomic polynomials prove that the sequence E'_p (with p a prime number) is $d\alpha/2\pi$-equidistributed. Prove the same result for E'_n.

2.2. EQUIDISTRIBUTION OF GALOIS ORBITS OF ALGEBRAIC POINTS

Let X be an algebraic projective variety defined over a number field K. For all field L containing K we denote by $X(L)$ the set of L rational points of X. Let $\overline{\mathbb{Q}}$ be the algebraic closure of $\mathbb{Q}$ and $\mathfrak{G}_K$ the Galois group of $\overline{\mathbb{Q}}$ over K. For all $x \in X(\overline{\mathbb{Q}})$ we define

$$E_x = \{x^\sigma \mid \sigma \in \mathfrak{G}_K\} \tag{2}$$

the Galois-orbit of x.

If we fix an embedding σ of K in $\mathbb{C}$, we can realize E_x as a subset of $X_\sigma(\mathbb{C})$. We write $\Delta_x = \mu_{E_x}$ the associated measure given by (1). A general (unsolved) problem is the following: let x_n be a sequence of points of $X(\overline{\mathbb{Q}})$, what can be said about the weak limits of the associated sequence $\Delta_n = \Delta_{x_n}$ of $\mathcal{P}(X)$.

We are not expecting a general answer to this question but we are going to give significant examples for which it's possible to say something. In all these examples there will be an underlying group structure on the variety X. To avoid some useless pathologies we make the following definition.

DEFINITION 2.5. Let X be an algebraic variety. A sequence x_n of points of X is said to be "generic" if for all proper algebraic subvariety $Y \subset X$ the set $\{n \in \mathbb{N} \mid x_n \in Y\}$ is finite.

(Exercise for topologists, prove that a sequence x_n is generic if an only if x_n converges to the generic point of X in the Zariski topology).

2.2.1. *The case of* $\mathbb{G}_m$: *a theorem of Bilu*

The first results (related to Exercise 2.4) are obtained by Bilu. Let $\mathbb{G}_m$ denote the multiplicative group, so $\mathbb{G}_m$ is an algebraic variety defined over $\mathbb{Q}$ and for all field K containing $\mathbb{Q}$ the set $\mathbb{G}_m(K)$ of K-rational points of $\mathbb{G}_m$ is K^*. There is a canonical height function

$$\hat{h}\colon \mathbb{G}_m(\overline{\mathbb{Q}}) \to \mathbb{R}_+ \tag{3}$$

satisfying the following conditions:

1. For all $\alpha \in \mathbb{G}_m(\overline{\mathbb{Q}})$ and all $n \in \mathbb{N}$, $\hat{h}(\alpha^n) = n\hat{h}(\alpha)$.
2. (Northcott) For all $n \in \mathbb{N}$ and all $X \in \mathbb{R}_+$ the set

$$\{\alpha \in \mathbb{G}_m(\overline{\mathbb{Q}}) \mid [\mathbb{Q}(\alpha)\colon \mathbb{Q}] \leq n \text{ and } \hat{h}(\alpha) \leq X\}$$

   is finite.

A first consequence of these properties is that $\hat{h}(\alpha) = 0$ if and only if α is a root of unity. In fact $\hat{h}(1) = \hat{h}(1^2) = 2\hat{h}(1) = 0$. If α is a root of unity then there exists a $n \in \mathbb{N}$ such that $\alpha^n = 1$. Hence $\hat{h}(\alpha^n) = n\hat{h}(\alpha) = \hat{h}(1) = 0$. If $\hat{h}(\alpha) = 0$ then for all $n \in \mathbb{N}$, $\hat{h}(\alpha^n) = 0$. Applying Northcott we find that $\{\alpha^n, n \in \mathbb{N}\}$ is a finite set; therefore α is a root of unity.

Another consequence of the Northcott's theorem is that if $\alpha_n \in \mathbb{G}_m(\overline{\mathbb{Q}})$ is a generic sequence of points such that $\hat{h}(\alpha_n) \to 0$ then $|E_{\alpha_n}| \to \infty$.

THEOREM 2.6 ((Bilu, 1997)). *Let $\alpha_n \in \mathbb{G}_m(\overline{\mathbb{Q}})$ be a generic sequence of points such that $\hat{h}(\alpha_n) \to 0$ then the sets E_{α_n} are $d\alpha/2\pi$-equidistributed.*

Proof. See (Bilu, 1997) for this statement and the generalization to the higher rank torus $\mathbb{G}_m^r$.

2.2.2. *The case of elliptic curves and Abelian varieties*

Let X be an elliptic curve over a field K, so X is an algebraic curve of genus 1 defined over K with the structure of an Abelian group on $X(K)$ (we denote by O the neutral element of $X(K)$). If $K = \mathbb{C}$ then X is isomorphic to $\Gamma\backslash\mathbb{C}$ for a lattice $\Gamma \subset \mathbb{C}$. The Lebesgue measure on $\mathbb{C}$ induces a canonical probability measure $\mu_X \in \mathcal{P}(X)$. From this description we see that the set $X[n] = \{P \in E(\mathbb{C}) \mid [n]P = O]\}$ is isomorphic to $\Gamma\backslash(1/n)\Gamma \simeq \mathbb{Z}/n\mathbb{Z} \oplus \mathbb{Z}/n\mathbb{Z}$.

EXERCISE 2.7. Prove that the sequences of subsets $X[n]$ of X is μ_X-equidistributed. (Use Proposition 2.2). Let $X'[n]$ be the subset of $X[n]$ consisting of points of order n (for example if p is prime $X'[p] = X[p] - \{O\}$). Prove that the sequence of subsets $X'[n]$ are μ-equidistributed.

Let X_K be an elliptic curve defined over a field K of characteristic 0. The group $\mathrm{End}(X_K)$ of K-endomorphism of X_K is $\mathbb{Z}$ or an order in an imaginary quadratic field. We say that X_K has complex multiplication if $\mathrm{End}(X_K) \neq \mathbb{Z}$.

Let X_K be a K-elliptic curve without complex multiplication. A consequence of Serre's open image theorem is that for all p big enough the Galois orbit of a point Q_p of order p is $X'[p]$. With the notation of the Section we have $E_{Q_p} = E'[p]$ and therefore the sets E_{Q_p} are μ_X-equidistributed. Using the full strength of Serre's open image theorem it's possible but not obvious to prove that if Q_n is a sequence of points of $X(\overline{\mathbb{Q}})$ with Q_n of order n then the sets E_{Q_n} are μ_X-equidistributed. If X_K has complex multiplication you may have infinitely many prime numbers p such that there exist a point Q_p of $X(\overline{\mathbb{Q}})$ of order p with E_{Q_p} of cardinality $p-1$ (corresponding to a cyclic isogenies of order p).

Using *Arakelov theory* it is possible to prove a very general result for the equidistribution of Galois-orbits of points with small heights on Abelian varieties (containing the equidistribution of the sets E_{Q_n} even for an elliptic curve with complex multiplication.)

An Abelian variety A of dimension g over $\mathbb{C}$ is a complex torus $A \simeq \Gamma\backslash\mathbb{C}^g$ endowed with the structure of a projective algebraic variety. The Lebesgue measure on $\mathbb{C}^n$ induces a canonical probability measure $\mu = \mu_A$ on A. We deduce from this description that $A(\mathbb{C})$ is an Abelian group. A point P of A is said to be a torsion point if there exists $n \in \mathbb{N}$ such that $[n]P = O$. The set

$$A[n] = \{P \in A(\mathbb{C}) \mid [n]P = 0\}$$

is isomorphic to $\mathbb{Z}/n\mathbb{Z}^{2g}$ as an Abelian group.

More generally an Abelian A_K variety over a field K is a projective algebraic variety endowed with the structure of an Abelian group structure. Concretely for all extension L of K $A(L)$ is an Abelian group. If K is a number field $A(K)$ is finitely generated (Mordell–Weil Theorem).

As in the case $\mathbb{G}_m$, if A_K is defined over a number field K it's possible to define a canonical height function

$$\hat{h}\colon A(\overline{\mathbb{Q}}) \to \mathbb{R}_+$$

(the Néron–Tate height) with the properties

1. For all $\alpha \in A(\overline{\mathbb{Q}})$ and all $n \in \mathbb{N}$, $\hat{h}([n]\alpha) = n^2\hat{h}(\alpha)$.
2. (Northcott) For all $n \in \mathbb{N}$ and all $X \in \mathbb{R}_+$ the set

$$\{\alpha \in A(\overline{\mathbb{Q}}) \mid [\mathbb{Q}(\alpha)\colon \mathbb{Q}] \leq n \text{ and } \hat{h}(\alpha) \leq X\}$$

is finite.

REMARK 2.8. The height function depends on the choice of a symmetric ample divisor. By definition of a projective variety we can find an ample line bundle $\mathcal{L}$ on A (sometime we say that $\mathcal{L}$ is a polarization). A line bundle $\mathcal{M}$ on A is said to be symmetric if $[-1]^*\mathcal{M} \simeq \mathcal{M}$. One can show that if $\mathcal{L}$ is ample then $[-1]^*\mathcal{L} \otimes \mathcal{L}$ is ample symmetric.

EXERCISE 2.9. Let A_K be an Abelian variety defined over a number field K. Prove that a point $P \in A(\overline{\mathbb{Q}})$ is torsion if and only if $\hat{h}(P) = 0$. If P_n is a generic sequence of points of $A(\overline{\mathbb{Q}})$ such that $\hat{h}(P) \to 0$ then the cardinality of the sets $E_{P_n} = \{P_n^\sigma \mid \sigma \in \mathfrak{G}_K\}$ tends to ∞. (*Hint*: try to imitate the case of $\mathbb{G}_m$). All these facts are independent of the choice made in defining the height.

The following results is due to Szpiro, Zhang and the author (Szpiro et al., 1997).

THEOREM 2.10. *Let A_K an Abelian variety defined over a number field K. For all embedding $\sigma: K \to \mathbb{C}$ we denote by μ_σ the canonical probability measure on $A_\sigma = A_K \otimes_\sigma \mathbb{C} \simeq \Gamma_\sigma \backslash \mathbb{C}^g$. Let P_n be a generic sequence of points of $A(\overline{\mathbb{Q}})$ such that $\hat{h}(P) \to 0$. Then for all $\sigma: K \to \mathbb{C}$ the sets $\sigma(E_{P_n})$ are μ_σ-equidistributed on A_σ.*

The proof uses Arakelov theory see (Szpiro et al., 1997).

2.2.3. *Equidistribution of* CM *elliptic curves*

The j-invariant establishes a bijection between $\mathbb{C}$ and the set of isomorphism classes of elliptic curves over $\mathbb{C}$. The endomorphism ring End(E) of an elliptic curve E over $\mathbb{C}$ is either $\mathbb{Z}$ or an order in an imaginary quadratic extension of $\mathbb{Q}$. An elliptic curve is said to be CM (meaning complex multiplication) if End(E) $\neq \mathbb{Z}$. A complex number x is said to be CM if the corresponding elliptic curve over $\mathbb{C}$ is CM.

Let us recall a few facts about CM elliptic curves. A CM elliptic curve is defined over $\overline{\mathbb{Q}}$. Let K be an imaginary quadratic extension of $\mathbb{Q}$ and $O_K \subset K$ be the ring of integers of K. Any order in O_K is of the form $O_{K,f} = \mathbb{Z} + fO_K$ for a unique integer $f \geq 1$. For $f \geq 1$ let $\Sigma_{K,f}$ be the set of isomorphism classes of pairs (E, α), with E a CM elliptic curve and $\alpha: O_{K,f} \to \mathrm{End}(E)$ an isomorphism of rings. The group $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$ acts transitively on $\Sigma_{K,f}$.

Let $\mathrm{Pic}(O_{K,f})$ be the Picard (or class) group of $O_{K,f}$ and $h = h_{K,f}$. Then the cardinality of $\Sigma_{K,f}$ is h. Let $H_{K,f}$ be the maximal Abelian extension of K which is unramified outside f. Then for all $E \in \Sigma_{K,f}$ we have $K(j(E)) = H_{K,f}$ and class field theory gives an isomorphism $\mathrm{Pic}(O_{K,f}) \simeq \mathrm{Gal}(H_{K,f}/K)$. Let $d_E = d_{K,f}$ be the absolute value of the discriminant of $O_{K,f}$. By the Siegel's theorem we get for all $\varepsilon > 0$

$$d_E^{1/2-\varepsilon} \underset{\varepsilon}{\ll} h = |\Sigma_{K,f}| \underset{\varepsilon}{\ll} d_E^{1/2} \log(d_E). \tag{4}$$

The modular group $\mathrm{SL}(2, \mathbb{Z})$ acts properly discontinuously on the upper half plane $\mathbb{H}$ by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} . z = \frac{az + b}{cz + d}.$$

The modular curve $Y = \mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H}$ is the set of isomorphism classes of elliptic curves over $\mathbb{C}$. The Poincaré measure $dx\,dy/y^2$ on $\mathbb{H}$ is $\mathrm{SL}(2,\mathbb{R})$-invariant and the volume of a fundamental domain for this measure is finite. (We therefore say that $\mathrm{SL}(2,\mathbb{Z})$ is a lattice of $\mathbb{H}$). Let $d\mu_0 = 3/\pi dx\,dy/y^2$ the induced probability measure on $Y \simeq \mathbb{C}$.

The following result is due do to Duke (Duke, 1988):

THEOREM 2.11 (Duke). *As $d_{K,f} \to \infty$ the $\Sigma_{K,f}$ are μ-equidistributed.*

The case of fundamentals discriminants (i.e. $f = 1$) is the main result of (Duke, 1988). The extension to the general case using Hecke operators is given in (Clozel and Ullmo, 2001, Theorem 2.4).

2.3. EQUIDISTRIBUTION OF HECKE POINTS

Let $\mathbb{H}$ be the upper half plane, $\Gamma = \mathrm{SL}(2,\mathbb{Z})$ and $Y = \Gamma\backslash\mathbb{H}$. Let $d\mu_0$ be the Poincaré metric and $D_0 = y^2(\partial^2/\partial x^2 + \partial^2/\partial y^2)$ the associated Laplacian. Let $L^2(\Gamma\backslash\mathbb{H}, d\mu_0)$ be the space of μ_0-square integrable Γ-invariant functions on $\mathbb{H}$.

The Hecke correspondences T_n on $X(1)$ are defined by

$$T_n.z = \sum_{ad=n}\sum_{0\leq b<d} \frac{az+b}{d}. \tag{5}$$

If $s \in \mathbb{C}$ we define $\sigma_s(n) = \sum_{d/n} d^s$. The degree of T_n as a correspondence is then $\sigma_1(n)$.

We have an induced action on functions on $X(1)$ given by:

$$\overline{T}_n f(z) = \frac{T_n}{\sigma_1(n)} f(z) = \frac{1}{\sigma_1(n)} \sum_{y\in T_n.z} f(y). \tag{6}$$

The following result is proved in (Clozel and Ullmo, 2001, Theorem 2.1). The proof is also taken from this paper. We just replaced the upper bound for eigenvalues of T_n given in (Bump et al., 1992) by the better bound obtained recently by Kim and Sarnak (Kim and Sarnak, 2003). (We therefore just replaced $5/28$ by $7/64$).

THEOREM 2.12.

(a) *For all f in $L^2(X(1), d\mu_0)$ $\overline{T}_n.f$ converges to $\int_{X(1)} f(\zeta)\, d\mu_0(\zeta)$ in $L^2(X(1), d\mu_0)$. More precisely For all f in $L^2(X(1), d\mu_0)$ and for all $\varepsilon > 0$, there exists a constant C_ε, (depending only on ε), such that*:

$$\left\|\overline{T}_n f - \int_{X(1)} f(\zeta)\, d\mu_0(\zeta)\right\| \leq C_\varepsilon n^{(-1/2)+(7/64)+\varepsilon} \|f\|. \tag{7}$$

(b) *Let f be a bounded C^∞ function on $X(1)$ such that $D_0 f$ is bounded. For all $\varepsilon > 0$, there exists $C_{\varepsilon,z,f}$ such that*

$$\left|\overline{T}_n f(z) - \int_{X(1)} f(\zeta)\, d\mu_0(\zeta)\right| \leq C_{\varepsilon,z,f} n^{(-1/2)+(7/64)+\varepsilon}. \tag{8}$$

(c) *For all bounded continuous function f on $X(1)$ and all $z \in X(1)$,we have*

$$\lim_{n\to+\infty} \overline{T}_n f(z) = \int_{X(1)} f(\zeta)\, d\mu_0(\zeta). \tag{9}$$

The convergence is uniform on compact sets.

2.3.1. *Spectral decomposition of $L^2(X(1), d\mu_0)$*
We have the “spectral decomposition”

$$L^2(X(1), d\mu_0) = \oplus_{n\geq 0}\mathbb{C}[\varphi_n] \oplus \mathcal{E} \tag{10}$$

with φ_n an orthonormal family of eigenfunctions of D_0 with associated eigenvalues $-\lambda_n$ and $\mathcal{E}$ is the continuous part of the spectrum. We write as usual s_n and r_n the complex numbers such that $\lambda_n = s_n(1 - s_n) = \frac{1}{4} + r_n^2$ (r_n is a real number). It’s possible to choose the φ_n eigenvectors for all the Hecke operators T_n. We suppose that this choice is made from now.

The part relative to the continuous spectrum is given by the following isometry:

$$E\colon L^2(\mathbb{R}_+) \to \mathcal{E} \tag{11}$$
$$h \mapsto \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} h(t) E_\infty\left(z, \frac{1}{2} + it\right) dt$$

here $L^2(\mathbb{R}_+)$ is the set of functions on $\mathbb{R}_+$ square integrable for the Lebesgue measure; $E_\infty(z, s)$ is the Eisenstein series at the cusp ∞, given by the formula

$$E_\infty(z, s) = \frac{1}{2} \sum_{(m,n)=1} \frac{1}{|mz + n|^{2s}}.$$

Let $\alpha \in L^2(X(1), d\mu_0)$ be spectrally decomposed as

$$\alpha(z) = \sum_{n\geq 0} A_n \varphi_n(z) + \int_0^{+\infty} h(t) E_\infty\left(z, \frac{1}{2} + it\right) dt. \tag{12}$$

Then

$$A_n = (\alpha, \varphi_n) = \int_X \alpha(z)\overline{\varphi}_n(z)\, d\mu_0(z)$$
$$h(t) = \frac{1}{2\pi} \int_X \alpha(z) E_\infty\left(z, \frac{1}{2} - it\right) d\mu_0(z)$$

(at least if α is C^∞ with compact support). The L^2 norm of α is then

$$||\alpha||^2 = \sum_n |A_n|^2 + 2\pi \int_0^{+\infty} |h(t)|^2 \, dt. \tag{13}$$

If $\alpha(z)$ is C^∞ with compact support then the spectral decomposition (12) is absolutely convergent and is uniformly convergent on compact sets (see (Iwaniec, 1995, Theorems 4.7 and 7.3)).

2.3.2. *Proof of Theorem* 2.12

We start the proof of Theorem 2.12 in the particular case $f = \varphi_n$ for some $n \in \mathbb{N}$ and the case where f appears in the continuous spectrum of D_0. The general case is then obtained using the spectral decomposition (12).

For $f = \varphi_0$ the constant function the Theorem 2.12 is clear as the degree of T_n as a correspondence is $\sigma_1(n)$.

LEMMA 2.13. *For all $k \geq 1$, and all $z \in X(1)$:*

$$\lim_{n\to\infty} \overline{T_n}\varphi_k(z) = 0 = \int_{X(1)} \varphi_k \, d\mu_0. \tag{14}$$

We know that φ_k is an eigenfunction of T_n. We define $\alpha_k(n)$ to be the associated eigenvalues:

$$T_n.\varphi_k = \alpha_k(n)\varphi_k.$$

The Ramanujan–Petersson conjecture predicts that:

$$|\alpha_k(n)| \leq d(n)n^{1/2},$$

where $d(n) = \sigma_0(n)$ is the number of positive divisors of n. The best known result towards the Ramanujan–Petersson conjecture is (Kim and Sarnak, 2003):

$$|\alpha_k(n)| \leq d(n)n^{(1/2)+(7/64)}. \tag{15}$$

This improves the result of (Bump et al., 1992) where the bound with $5/28$ instead of $7/64$ was obtained.

For all $\varepsilon > 0$ and n big enough

$$|\overline{T_n}\varphi_k(z)| = \left|\frac{\alpha_k(n)\varphi_k(z)}{\sigma_1(n)}\right| \leq n^{(-1/2)+(7/64)+\varepsilon}|\varphi_k(z)|. \tag{16}$$

LEMMA 2.14. *Let f be a function in $\mathcal{E} \cap D(X(1))$. Then*

$$\lim_{n\to\infty} \overline{T_n}f(z) = 0 = \int_{X(1)} f \, d\mu_0. \tag{17}$$

Proof. There exists a function $h(t) \in L^2(\mathbb{R}_+)$ such that:

$$f(z) = \int_0^\infty h(t)E_\infty(z, \tfrac{1}{2} + it)\,dt$$

This last integral is absolutely convergent. We recall that $E_\infty(z, s)$ is an eigenform of the Hecke operators T_n:

$$T_nE_\infty(z, s) = n^s\sigma_{1-2s}(n)E_\infty(z, s).$$

We therefore obtain

$$T_nf(z) = n^{1/2}\int_0^\infty n^{it}\sigma_{-2it}(n)h(t)E_\infty(z, \tfrac{1}{2} + it)\,dt. \tag{18}$$

Therefore for all $\varepsilon > 0$ and all $n \gg 0$ we get:

$$|\overline{T}_nf(z)| \leq n^{(-1/2)+\varepsilon}\int_0^\infty |h(t)E_\infty(z, s)|\,dt. \tag{19}$$

We can now give the proof of Theorem 2.12. Let f be a function in $L^2(X(1), d\mu_0)$. The spectral decomposition of f is written:

$$f(z) = \sum_{k\geq 0} A_k\varphi_k(z) + \int_0^{+\infty} h(t)E_\infty(z, \tfrac{1}{2} + it)\,dt. \tag{20}$$

We define

$$J_n = \left\| \overline{T}_nf - \int_{X(1)} f(\zeta)\,d\mu_0(\zeta) \right\|.$$

Then

$$J_n = \left\| \sum_{k\geq 1} \frac{A_k\alpha_k(n)\phi_k}{\sigma_1(n)} + \frac{n^{1/2}}{\sigma_1(n)}\int_0^\infty n^{it}\sigma_{-2it}(n)h(t)E_\infty(z, s)\,dt \right\|.$$

Using (13), we obtain:

$$J_n^2 = \frac{1}{\sigma_1(n)^2}\sum_{k\geq 1} |A_k|^2|\alpha_k(n)|^2 + 2\pi\frac{n}{\sigma_1(n)^2}\int_0^\infty |h(t)|^2|\sigma_{-2it}(n)|^2\,dt.$$

The proof of the (a) of (2.12) is obtained using the upper bounds for $\alpha_k(n)$ given in the equation (15).

We define $D(X(1))$ as the space of C^∞ bounded functions on $X(1)$ such that D_0f is bounded.

Let $f \in D(X(1))$ and $z \in X(1)$. Using (16) and (19) we find for all $\varepsilon > 0$ and all $n \gg 0$ the upper bound:

$$|\overline{T}_n f(z) - A_0| \leq n^{(-1/2)+(5/28)+\varepsilon} \left(\sum_{k \geq 1} |A_k \varphi_k(z)| + \int_0^\infty |h(t) E_\infty(z, s)|\, dt \right).$$

This ends the proof of the part (b) of (2.12) as $A_0 = \int_{X(1)} f(z)\, d\mu_0(z)$ and as $f \in D(X(1))$ the spectral decomposition is absolutely convergent

We finally suppose that $f \in C_0(X(1))$. Let $z \in X(1)$ and $\varepsilon > 0$. We can find $\phi \in D(X(1))$ such that

$$\sup_{x \in X(1)} |f(x) - \phi(x)| \leq \varepsilon.$$

Using part (b) of the Theorem 2.12, we know that for all $n \gg 0$:

$$\left| \overline{T}_n \phi(z) - \int_{X(1)} \phi\, d\mu_0 \right| \leq \varepsilon.$$

We define l_n as $I_n = |\overline{T}_n f(z) - \int_{X(1)} f\, d\mu_0|$. We therefore obtain that

$$I_n \leq |\overline{T}_n f(z) - \overline{T}_n \phi(z)| + \left| \overline{T}_n \phi(z) - \int_{X(1)} \phi\, d\mu_0 \right| + \left| \int_{X(1)} (\phi - f)\, d\mu_0 \right| \leq 3\varepsilon.$$

This ends the proof of the part (c) of the Theorem 2.12.

2.3.3. *Higher rank generalization*

The result of the previous sections can be generalized to an arbitrary almost simple simply connected linear group $G_{\mathbb{Q}}$ (as $\mathrm{SL}(n)_{\mathbb{Q}}$ or $\mathrm{Sp}(n)_{\mathbb{Q}}$). A proof using harmonical analysis as in the previous part is given in (Clozel et al., 2001; Clozel and Ullmo, 2001). The method gives a convergence rate which is often optimal. This is the case for $\mathrm{SL}(n)_{\mathbb{Q}}$ or $\mathrm{Sp}(n)_{\mathbb{Q}}$ if $n \geq 3$. Note that optimal results are obtained without using the (unknown) generalized Ramanujan conjecture for parameters of automorphic representation of $\mathrm{SL}(n)_{\mathbb{Q}}$ or $\mathrm{Sp}(n)_{\mathbb{Q}}$. The extension to arbitrary reductive groups is in general easy. A proof of a slightly more general result (without a convergence rate) is obtained by Eskin and Oh (Eskin and Oh, 2006) by ergodic methods.

3. The Manin–Mumford and the André–Oort Conjecture

3.1. ABSTRACT FORM OF THE CONJECTURES

The Manin–Mumford conjecture about torsion points of Abelian varieties and the André–Oort conjecture about CM points on Shimura varieties (Ex.: the moduli space of principally polarized Abelian varieties, Hilbert modular varieties or product of modular curves...) can abstractly be stated in a unified way. The purpose of this section is to explain these conjectures and the relation with some theorems or conjectures about equidistribution.

Let X be an algebraic variety over $\mathbb{C}$. Let $\mathcal{S}(X)$ be a set of irreducible subvarieties of X. A subvariety $Z \in \mathcal{S}(X)$ is called special and a special subvariety of dimension 0 is called a special point. We say that $\mathcal{S}(X)$ is an admissible set of special subvarieties if:

1. $X \in \mathcal{S}(X)$.
2. For all $Z \in \mathcal{S}(X)$ the set of special points $x \in Z$ is Zariski dense in Z.
3. An irreducible component of an intersection of special varieties is a special variety.

REMARK 3.1. As a consequence of property 3, if W is a subset of $X(\mathbb{C})$ there exists a smallest special subvariety Z_W among special subvarieties containing W.

The main examples of admissible sets of special subvarieties are:

(i) An Abelian variety $X = A$, $\mathcal{S}(A)$ is the set of torsion subvarieties. A torsion subvariety is the translate by a torsion point of an Abelian subvariety. The special points are the torsion points.

(ii) A torus $X = T$, $\mathcal{S}(T)$ is the set of torsion subvarieties. A torsion subvariety is the product of a point of finite order by a subtorus. The special points are the points of finite order.

(iii) A Shimura variety $X = S$, $\mathcal{S}(S)$ is the set of subvarieties of Hodge type. A subvariety of Hodge type is an irreducible component of the translate by a Hecke operator of a sub-Shimura variety. The special points are the CM points. This case will be detailed in Section 3.3.

Note as a general rule that everything is known in the case of an Abelian variety or in the case of a torus but despite recent progress the case of Shimura variety is mainly conjectural. Other situations as mixed Shimura varieties (see (Pink, 2005)) or semi-Abelian varieties (see (Chambert-Loir, 2000; David and Philippon, 2000)) can be considered. It's possible that other situations coming from variations of Hodge structures could be considered.

CONJECTURE 3.2 (Abstract form). *There are* 2 *equivalent ways of formulating the conjecture*:

(a) *An irreducible component of the Zariski closure of a set of special points is a special subvariety.*

(b) *Let Y be an algebraic subvariety of X. There exists special subvarieties $\{Z_1, \dots, Z_r\}$ with $Z_i \subset Y$ such that if $Z \subset Y$ is a special subvariety then*

$$Z \subset \bigcup_{i=1}^{r} Z_i.$$

The conjecture in this abstract way is certainly too optimistic. For example you could take for X any projective variety of dimension $g \geq 2$ and for $\mathcal{S}(X)$ the union of X and the set of all points of X. (I don't know such a trivial counterexample if we impose that $\mathcal{S}(X)$ is countable in the definition of an admissible set). Nevertheless it may be useful to understand it in this form to see what is really used in the important examples. Note that if $Y \subset X$ is a curve the conjecture predicts that Y is special if and only if Y contains infinitely many special points. Let's prove that the two forms of the conjecture are indeed equivalent:

Let Y be an algebraic subvariety of X and Σ_Y the set of special points contained in Y. Let $\{Z_1, \dots, Z_r\}$ be the components of the Zariski closure of Σ_Y. If (a) is true then the Z_i are special and have the properties of (b).

Let Σ be a set of special points and Y a component of the Zariski closure. By (b) there exists a finite set $\{Z_1, \dots, Z_r\}$ of special subvarieties of Y such that all the special subvarieties of Y are contained in one of the Z_i. As Y is the Zariski closure of Σ, $Y \subset \bigcup_{i=1}^{r} Z_i$ and there exists $i \in \{1, \dots, r\}$ such that $Y = Z_i$. Therefore Y is special.

The theory is even more interesting when:

(a) The variety X is defined over a number field K and the special points are defined over $\overline{\mathbb{Q}}$.

(b) A special subvariety Z of X is canonically endowed with a probability measure μ_Z such that the Zariski closure of $\mathrm{Supp}(\mu_Z)$ is Z.

DEFINITION 3.3. An admissible set $\mathcal{S}(X)$ of special subvarieties of X with properties (a) and (b) is said to be strongly admissible.

The property (a) implies that the special subvarieties of X are defined over number field. (A subvariety containing a dense set of points defined over $\overline{\mathbb{Q}}$ is defined over $\overline{\mathbb{Q}}$). If P is a special point the canonical probability measure on P is $\mu_P = \delta_P$. As in Section 2.2 we fix an embedding of K in $\mathbb{C}$ and $X(\overline{\mathbb{Q}})$ is realized as a subset of $X(\mathbb{C}) = X$. Let $E_P = E_{P,K} = \{P^\sigma, \sigma \in \mathfrak{G}_K\}$ and

$$\Delta_P = \Delta_{P,K} = \frac{1}{|E_P|} \sum_{x \in E_P} \delta_y.$$

DEFINITION 3.4. Let X be a variety and $\mathcal{S}(X)$ a strongly admissible set of special subvarieties. A sequence P_n of points in $X(\mathbb{C})$ is said to be *strict* (relatively to $(X, \mathcal{S}(X))$ if for all special subvariety $Z \neq X$ of X the set $\{n \in \mathbb{N} \mid P_n \in Z\}$ is finite.

The expected equidistribution conjecture is

CONJECTURE 3.5 (Abstract form). *Let X be a variety and $\mathcal{S}(X)$ an admissible set of special subvarieties. Let K be a number field over which X is defined. Let P_n be a strict sequence of special points of $X(\mathbb{C})$ then the sets $E_{P_n,K}$ are μ_X-equidistributed: the associated sequence of probability measure $\Delta_{n,K} = \Delta_{P_n,K}$ weakly converges to μ_X.*

PROPOSITION 3.6. *Conjecture* 3.5 *implies Conjecture* 3.2.

Let Σ be a set of special points and Y a component of the Zariski closure of Σ. Then $\Sigma_Y = \Sigma \cap Y$ is a Zariski dense subset of special points of Y. Let $Z = Z_Y$ be the smallest special subvariety of X containing Y. The set $\mathcal{S}(Z)$ of special subvarieties of X contained in Z is strongly admissible. The subvariety Z is defined over a number field L.

LEMMA 3.7. *There exists a strict sequence of special points of* Σ_Y *relatively to* $(Z, \mathcal{S}(Z))$.

The set of special subvarieties is countable as special subvarieties are defined over $\overline{\mathbb{Q}}$. We can therefore write $\mathcal{S}(Z) = \{(Z_n), n \in \mathbb{N}\}$. For all $n \in \mathbb{N}$ we define

$$\Sigma_{n,Y} = \left\{ P \in \Sigma_Y \,\middle|\, P \notin \bigcup_{i=1}^{n} Z_i \right\}.$$

As Σ_Y is Zariski dense in Y, for all $n \in \mathbb{N}$, $I_n \neq \emptyset$. We can therefore choose $P_n \in \Sigma_{n,Y}$. By construction P_n is a strict sequence.

Using Conjecture 3.5 we see that the sequence $\Delta_{P_n,L}$ weakly converges to μ_Z. As $\mathrm{Supp}(\Delta_{P_n,L})$ is contained in L for all n (and Y is closed) we find that $\mathrm{Supp}(\mu_Z) \subset Y$. As the Zariski closure of $\mathrm{Supp}(\mu_Z)$ is Z (by property (b)), $Y = Z$. Therefore Y is a special subvariety as predicted by Conjecture 3.2.

In fact a even more general result is expected. Let X be an algebraic variety defined over a number field K and $\mathcal{S}(X)$ a strongly admissible set of special subvarieties. For all $Z \in \mathcal{S}(X)$, the set

$$O(Z) = \{Z^{\sigma} \mid \sigma \in \mathfrak{G}_K\}$$

is finite and contained in $\mathcal{S}(X)$. Let Δ_Z be the measure

$$\Delta_Z = \frac{1}{|O(Z)|} \sum \mu_{Z_{\sigma}}.$$

Let $\mathcal{P}(X)$ be the set of Borel probability measure on X and

$$\mathcal{Q}(X) = \{\Delta_Z \mid Z \in \mathcal{S}(X)\}.$$

The most optimistic conjecture about equidistribution is:

CONJECTURE 3.8. *The subset $\mathcal{Q}(X)$ of $\mathcal{P}(X)$ is compact. If Δ_{Z_n} is a sequence of measure in $\mathcal{Q}(X)$ weakly converging to μ_Z then for all $n \gg 0$,* $\mathrm{Supp}(\mu_{Z_n}) \subset \mathrm{Supp}(\mu_Z)$.

We will discuss results for this conjecture for Abelian varieties in Sections 3.2 and some related results for some sequences μ_{Z_n} where Z_n is a special subvariety (and therefore geometrically irreducible) in Section 4.

3.2. THE MANIN–MUMFORD AND THE BOGOMOLOV CONJECTURE

In the case of Abelian varieties, all the abstract theory of the previous section is proved. If A is an Abelian variety, a special point is a torsion point and a special variety is a torsion variety. Let $\mathrm{Tor}(A)$ be the set of torsion point of A. The Conjecture 3.2 in this case is due to Manin and Mumford:

THEOREM 3.9. *Let A be an Abelian variety and Σ and X a subvariety of A. Then*

$$X \cap \mathrm{Tor}(A) = \bigcup_{i=1}^{r} T_i \cap \mathrm{Tor}(A)$$

for some torsion subvarieties $(T_1, \dots, T_r)$.

A first proof of the conjecture was given by Raynaud (Raynaud, 1983), (see (Raynaud, 1988) for the case of a curve) using p-adic method. Hindry (Hindry, 1988) gave a proof using Galois theory and Diophantine approximation. Hrushowski (Hrushowski, 2001) gave a proof using ideas from logic (model theory of field). As model theory of field is not so far from the theory of constructible set in algebraic geometry it's not completely a surprise that Pink and Roesler (Pink and Roessler, 2004) where able to translate in a short and efficient way Hrushowski's proof in the language of algebraic geometry (and some Galois theory). Finally a proof using Arakelov theory and ideas from "equidistribution of points with small height" of the Bogomolov conjecture (to be discussed later in this section) was given by Zhang (Zhang, 1998) and the author (Ullmo, 1998). As the Bogomolov conjecture contains the Manin–Mumford conjecture, this gives an almost completely analytic proof of the Manin–Mumford conjecture.

Recall that if $(a_n)_{n\in\mathbb{N}}$ is a sequence of algebraic points of an Abelian variety A defined over a number field we say that (a_n) is a generic sequence (resp. a strict sequence) If for any proper subvariety $Y \subset A$ (resp. for any proper torsion subvariety $Y \subset A$) the set

$$\{n \in \mathbb{N}, a_n \in A(\overline{\mathbb{Q}})\}$$

is finite.

REMARK 3.10. With these definitions we can rephrase the Manin–Mumford conjecture in the following way: "any strict sequence of torsion point of $A(\overline{\mathbb{Q}})$ is generic". If a_n is a strict sequence of torsion point of $A(\overline{\mathbb{Q}})$, Y a proper subvariety of A such that $T_Y = \{n \in \mathbb{N}, a_n \in Y(\overline{\mathbb{Q}})\}$ is not finite. The Manin–Mumford conjecture implies that the components of the Zariski closure of the a_n with $n \in T_Y$ are torsion subvarieties containing infinitely many terms of the sequence a_n. This contradicts the hypothesis that a_n is strict. The other direction can be proved in the same lines as the Proposition 3.6 and is left as an exercise.

When you combine the Manin–Mumford conjecture and the Theorem 2.10 you obtain the following results (Szpiro et al., 1997) in the direction of the Conjecture 3.8:

THEOREM 3.11. *Let A_K an Abelian variety defined over a number field K. For all embedding $\sigma\colon K \to \mathbb{C}$ we denote by μ_σ the canonical probability measure on $A_\sigma = A_K \otimes_\sigma \mathbb{C} \simeq \Gamma_\sigma\backslash\mathbb{C}^g$. Let P_n be a strict sequence of torsion point of $A(\overline{\mathbb{Q}})$. Then for all $\sigma\colon K \to \mathbb{C}$ the sets $\sigma(E_{P_n})$ are μ_σ-equidistributed on A_σ.*

For Abelian varieties, the full Conjecture 3.8 is a consequence of the extension of this last result to the equidistribution of Galois orbits of special subvarieties due (independently) to Autissier (Autissier, 2004) and Baker–Ih (Baker and Ih, 2004).

The Bogomolov conjecture is a generalization of the Manin–Mumford conjecture once we recall that a point P of an Abelian variety defined over a number field is a torsion point if and only if the Néron–Tate heights $\hat{h}(P)$ of P is 0 :

CONJECTURE 3.12 (Bogomolov). *Let A be an Abelian variety defined over a number field. Let Y be a non-torsion subvariety of A. There exists $c > 0$ such that the set*

$$\{P \in Y(\overline{\mathbb{Q}}) \mid \hat{h}(P) < c\}$$

is not Zariski-dense in Y.

The idea behind this conjecture is the following. Lang's conjecture predicts that the set of *rational points* $Y(K)$ of a variety of general type over a number field K should not be Zariski dense in Y. This has been checked by Faltings (Faltings, 1991) for non-torsion varieties (the case of curve (Faltings, 1984b) is the celebrated Mordell conjecture). Such a variety certainly contains infinitely many *algebraic points* but $Y(\overline{\mathbb{Q}})$ is not too big: it's a discrete set in the Néron–Tate topology.

As in the Remark 3.10, the Bogomolov conjecture is equivalent to the statement that "any strict sequences a_n of points of $A(\overline{\mathbb{Q}})$ such that $\hat{h}(a_n) \to 0$ is a generic sequence". The statement of Theorem 3.11 remains true with "torsion point" replaced by points with Néron–Tate height tending to 0.

The proof of this conjecture in the case of a curve in its Jacobian is given in (Ullmo, 1998) and the general case is proved along the same lines in (Zhang, 1998). It's unfortunately beyond the scope of these notes to give a detailed account of the proof of the Bogomolov conjecture. The interested reader can read the account given in Bourbaki's seminar by Abbes (Abbes, 1997).

Let's just sketch the principle of the proof in the case of a curve in its Jacobian. The starting point is a general theorem about the "equidistribution of generic sequences of points with small heights" (Szpiro et al., 1997) for more general heights than the Néron–Tate height on Abelian varieties.

Let X a curve of genus $g \geq 2$ defined over a number field K and fix an embedding ϕ of X in its Jacobian J. The canonical height $\hat{h}$ on $J(\overline{\mathbb{Q}})$ induces a canonical height on $X(\overline{\mathbb{Q}})$. Fix an embedding of K in $\mathbb{C}$, then $X_{\mathbb{C}}$ is a Riemann surface. We have a natural Hermitian inner product on the space $H^0(X_{\mathbb{C}}, \Omega^1_X)$ of holomorphic differential forms on $X_{\mathbb{C}}$ given by

$$(\alpha,\beta) = \frac{i}{2} \sum_X \alpha \wedge \bar{\beta}.$$

Let $\{\omega_1, \dots, \omega_g\}$ be an orthonormal basis of $H^0(X_{\mathbb{C}}, \Omega^1_X)$. Then we define a canonical $(1, 1)$-form μ on $X_{\mathbb{C}}$ by setting

$$\mu := \frac{i}{2g} \sum_{k=1}^{g} \omega_k \wedge \overline{\omega_k}.$$

The form μ does not depend on a choice of an orthonormal basis. The associated measure μ is called the canonical or the Arakelov measure.

Let $P_n \in X_K(\overline{\mathbb{Q}})$ be a generic sequence such that $\hat{h}(P_n) \to 0$. The result of (Szpiro et al., 1997) implies that the associated sequence of Galois orbits (as defined in Section 2.2) converges weakly to μ.

Let ϕ_g: $jX^g \to J$ be the morphism $(x_1, \dots, x_g) \mapsto \sum_{i=1}^{g} \phi(x_i)$. Let $\pi_i = X^g \to X$. By a diagonal process, it's possible to construct a generic sequence $y_n = (x_{1,n}, \dots, x_{g,n})$ of $X^g(\overline{\mathbb{Q}})$ such that for all i, $\hat{h}(x_{i,n}) \to 0$. Using the result of (Szpiro et al., 1997), we find that the associated sequence of Galois orbits converges weakly to the measure

$$\mu_g = \pi_1^* \mu \wedge \cdots \wedge \pi_g^* \mu.$$

As $z_n = \phi_g(y_n)$ is a generic sequence of $J(\overline{\mathbb{Q}})$ such that $\hat{h}(z_n) \to 0$, using Theorem 2.10 we know that the associated sequence of Galois orbits converges weakly to the normalized Haar measure μ_J of J.

Combining the two results (and using easy results about the morphism ϕ_g) we obtain the equality:

$$\phi_g^* \mu_J = g! \mu_g = g! \pi_1^* \mu \wedge \cdots \wedge \pi_g^* \mu.$$

It's easy to see that μ_g is everywhere positive and that $\phi_g^* \mu_J$ is 0 at the points where the morphism ϕ_g is singular (for example at $(P_0, \dots, P_0)$ for a Weierstrass point of $X_{\mathbb{C}}$). This contradiction finishes the proof.

3.3. THE ANDRÉ–OORT CONJECTURE

The André–Oort conjecture is the analogue for Shimura varieties of the Manin–Mumford conjecture for Abelian varieties. It's not possible to give here a complete account of Shimura varieties. The interested reader should see (Deligne, 1971; Deligne, 1979; Milne, 2006) but two aspects should be kept in mind.

1. Shimura varieties are Hermitian locally symmetric spaces.
2. Shimura varieties are moduli-spaces for interesting objects as Abelian varieties.

The aim of this part is to describe the special points and special subvarieties in this context and to formulate the André–Oort conjecture. We will focus on examples.

Hermitian locally symmetric space. Let $G = G_{\mathbb{Q}}$ be a connected reductive group over $\mathbb{Q}$, $G(\mathbb{R})^+$ the connected component of 1 of $G(\mathbb{R})$ and K_∞ a maximal compact subgroup of $G(\mathbb{R})$. Let $Z(G)$ be the center of G. Then G is the almost direct product

$$G \simeq Z(G)G_1G_2 \cdots G_r$$

for some $\mathbb{Q}$-simple groups G_i. We make the following assumption:

(*): For all $i \in \{1, \dots, r\}$, $G_i(\mathbb{R})$ is not compact.

The space $X^+ = G(\mathbb{R})/Z(G)(\mathbb{R})K_\infty$ is called a symmetric space. When X^+ is endowed with an $H(\mathbb{R})^+$-invariant complex structure we say that X^+ is an Hermitian symmetric space. A couple $(G_{\mathbb{Q}}, X^+)$ is called a (connected) *Shimura datum.* Deligne (Deligne, 1971; Deligne, 1979) proved that such an X^+ is a connected component of the $G(\mathbb{R})$-conjugacy class X of a morphism of algebraic groups

$$\alpha \colon \mathbb{S} \colon \to G_{\mathbb{R}}$$

Here $\mathbb{S} = \mathrm{Res}_{\mathbb{C}/\mathbb{R}}\mathbb{G}_m$ is the Deligne torus (so $\mathbb{S}$ is $\mathbb{C}^*$ as an algebraic group.) If $x \in X^+$, we'll write $x(\mathbb{S}) \subset G(\mathbb{R})$ for the image of the associated morphism $x \colon \mathbb{S} \to G_{\mathbb{R}}$. A Shimura datum is defined in (Deligne, 1971) as a couple $(G_{\mathbb{Q}}, X)$.

The simple groups $G_{\mathbb{R}}$ such that $X^+ = G(\mathbb{R})/K_\infty$ is Hermitian symmetric are well known inside the classification of linear semi-simple non-compact groups over $\mathbb{R}$. For example the symplectic group $\mathrm{Sp}(2, g)$, unitary groups $U(p, q)$ or orthogonal groups $\mathrm{So}(N, 2)$ have an associated symmetric space which is Hermitian. The symmetric spaces associated to $\mathrm{SL}(n, \mathbb{R})$ ($n \geq 3$) or $\mathrm{So}(p, q)$ (with $p \neq 2$ and $q \neq 2$) are not Hermitian.

A subgroup Γ of $G(\mathbb{Q})^+ = G(\mathbb{Q}) \cap G(\mathbb{R})^+$ is called an arithmetic lattice if Γ is commensurable to $G_{\mathbb{Z}}(\mathbb{Z})$ for a $\mathbb{Z}$-structure on $G_{\mathbb{Q}}$. This notion is independent of a choice of a $\mathbb{Z}$-structure on $G_{\mathbb{Q}}$. A standard way of producing such a $\mathbb{Z}$-structure is to fix an embedding of $G_{\mathbb{Q}}$ in $\mathrm{GL}(n, \mathbb{Q})$ and to take for $G_{\mathbb{Z}}$ the Zariski closure of $G_{\mathbb{Q}}$ in $\mathrm{GL}(n, \mathbb{Z})$.

Any symmetric space $X = G(\mathbb{R})/K_\infty$ is endowed with a $G(\mathbb{R})$-invariant measure. If Γ is an arithmetic lattice, this measure induces a measure on $\Gamma \backslash X^+$. The volume of $S = \Gamma \backslash X^+$ is finite for this measure (hence the notion of lattice—see Borel (Borel, 1969) for a proof). A space of the form $\Gamma \backslash X^+$ for a lattice Γ is called a locally symmetric space. Any locally symmetric space S is therefore endowed with a canonical probability measure μ_S.

If X^+ is Hermitian symmetric and Γ is an arithmetic lattice then $S = \Gamma\backslash X^+$ is endowed with a complex structure. Such a S is an (arithmetic) Hermitian locally symmetric space. The main fact is the relation with the world of algebraic geometry:

Baily-Borel. There exists a unique structure of algebraic variety on $S = \Gamma\backslash X^+$ over $\mathbb{C}$ such that for any algebraic variety T, any analytic morphism from T to S is induced from a morphism of algebraic varieties. With this structure S is quasi-projective. If Γ is torsion free then S is smooth.

If moreover $\Gamma \subset G(\mathbb{Q})$ is a congruence lattice (Γ contains the Kernel $\Gamma(N)$ of the map $G_{\mathbb{Z}}(\mathbb{Z}) \to G_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z})$ for some $n \in \mathbb{N}$) we say that $S = \Gamma\backslash X^+$ is a "connected Shimura variety".

EXAMPLE 3.13. $G = \mathrm{SL}(2, \mathbb{Q})$, $K_\infty = \mathrm{SO}_2(\mathbb{R})$, $X^+ = \mathbb{H}$ is the upper half plane. If Γ is a congruence subgroup of $\mathrm{SL}(2, \mathbb{Q})$ $S = \Gamma\backslash\mathbb{H}$ is a modular curve and S is a moduli space for elliptic curves with an additional structure defined by Γ.

EXAMPLE 3.14. $G = GS_p(2g, \mathbb{Q})$, $K_\infty = U_g(\mathbb{C})$, $X^+ = \mathbb{H}_g$ is the Siegel–Half plane. If $\Gamma = \Gamma(N)$ then $S = \mathcal{A}_{g,N}$ is the moduli space of principally polarized Abelian varieties of dimension g with full N-level structure.

EXAMPLE 3.15. Let F be a totally real extension of $\mathbb{Q}$ of degree g. $G_{\mathbb{Q}} = \mathrm{Res}_{F/\mathbb{Q}}\,\mathrm{SL}(2, F)$. Then $G(\mathbb{R}) = \mathrm{SL}(2, \mathbb{R})^g$, $K_\infty = \mathrm{SO}_2(\mathbb{R})^g$, $X^+ = \mathbb{H}^g$. Take $\Gamma = \mathrm{SL}_2(O_F)$, then $S = \mathrm{SL}_2(O_F)\backslash\mathbb{H}^g$ is an Hilbert modular variety parametrizing polarized Abelian varieties A of dimension g with an imbedding of O_F in the endomorphism of A.

EXAMPLE 3.16. Let F be as before and B a quaternion algebra over F. Then $B(\mathbb{R}) = M_2(\mathbb{R})^d \times \mathbf{H}^{g-d}$ (where $\mathbf{H}$ is the usual quaternions over $\mathbb{R}$). Let $G_{\mathbb{Q}}$ be the group of elements of B^* with reduced norm 1 and Γ the units of norm one in B^*. Then $G(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})^d \times \mathrm{SO}_3(\mathbb{R})^{g-d}$, $K_\infty = \mathrm{SO}_2(\mathbb{R})^d \times \mathrm{SO}_3(\mathbb{R})^{g-d}$ and $X^+ = \mathbb{H}^d$. Then $S = \Gamma\backslash\mathbb{H}^d$ is a "quaternionic Shimura" variety. Example 3.15 corresponds to $B = M(2, F)$. When $d = 1$, S is a curve (a Shimura curve). Any curve which is a "connected Shimura" variety is obtained from such a quaternion algebra.

Hecke correspondences. Let $S = \Gamma\backslash G(\mathbb{R})^+/K_\infty = \Gamma\backslash X^+$ a connected Shimura variety. Let $q \in G(\mathbb{Q})$, as Γ is an arithmetic lattice $q^{-1}\Gamma q$ is commensurable to Γ: $\Gamma \cap q^{-1} \cap \Gamma q$ is of finite index in Γ and $q^{-1}\Gamma q$. Let $C(\Gamma)$ be the commensurator of Γ

$$C(\Gamma) = \{g \in G(\mathbb{R}) \mid g\Gamma g^{-1} \text{ commensurable with } \Gamma\}.$$

If $G = G^{ad}$ then $C(\Gamma) = G(\mathbb{Q})$, for a general reductive group over $\mathbb{Q}$ (see (Platonov and Rapinchuk, 1994, Proposition 4.6, p. 206)), Let q be an element of $C(\Gamma)$. Let $S_q = \Gamma \cap q^{-1}\Gamma q \backslash X^+$ and α_q the finite map $S_q \to S$ induced from the inclusion $\Gamma \cap q^{-1} \cap \Gamma q \subset \Gamma$. The translation by q on X^+ (given by $x \mapsto g.x$ induces a second finite morphism $\beta_q \colon S_q \to S$.

Let T_q be the image in $S \times S$ of S_q by the map (α_q, β_q). Then T_q is an algebraic correspondence on S. Such a correspondence is called a modular correspondence. For all $x \in S$, we have "$T_q.s = \beta_q(\alpha_q^{-1}(x))$" where we have to count with multiplicities the points in $(\alpha_q^{-1}(x))$. If Γ acts on X^+ without fixed points the maps α_q and β_q are unramified (therefore for all $x \in S$, $(\alpha_q^{-1}(x))$ has exactly $\deg(\alpha) = [\Gamma \cap q^{-1} \cap \Gamma q \colon \Gamma]$ points). If $S = \mathrm{SL}(2, \mathbb{Z}) \backslash \mathbb{H}$ and $q = \begin{pmatrix} l & 0 \\ 0 & 1 \end{pmatrix}$ for a prime number l then T_q is the usual Hecke operator T_l discussed in Section 2.3.

Special subvarieties. Let S be a connected Shimura variety, we would like to define a set $\mathcal{S}(S)$ of special subvarieties with the properties $(1, 2, 3)$ of Section 3.1. A subvariety Z in $\mathcal{S}(S)$ should be (using our two points of view) a sub-Hermitian locally symmetric space and a moduli space for objects of S with some additional structures (as polarization, level, endomorphism. . .).

If (G_1, X_1^+) is a Shimura datum with $G_1 \subset G$ (as $\mathbb{Q}$-algebraic groups) inducing an inclusion $X_1^+ \subset X$, we say that (G_1, X_1^+) is a sub-Shimura datum. $\Gamma_1 = \Gamma \cap G_1(\mathbb{R})^+$ is an arithmetic lattice of G_1. A special subvariety of S is the image S_1 of $\Gamma_1 \backslash X_1^+$ for a sub-Shimura-datum (G_1, X_1^+).

For all $x \in X^+$ the subgroup of $G(\mathbb{R})^+$ fixing x is the product of $Z(G)(\mathbb{R}) \cap G(\mathbb{R})^+$ and a maximal compact subgroup K_x of $G(\mathbb{R})^+$. If there exists a torus $T_{\mathbb{Q}}$ of $G_{\mathbb{Q}}$ such that

$$x(\mathbb{S}) \subset T(\mathbb{R})^+ \subset Z(G)(\mathbb{R})^+ K_x$$

then $(T_{\mathbb{Q}}, \{x\})$ is a sub-Shimura datum. We say that x is a special point of X^+ and its image in S is a special point of S. The set of special points of S is obtained in this way.

The relation with the theory of complex multiplication is the following. A CM-field is a totally imaginary extension of degree 2 of a totally real number field. A simple Abelian variety A of dimension g is CM if the endomorphism $\mathrm{End} A \otimes \mathbb{Q}$ of A is a CM field of dimension $2g$. An Abelian A variety is said to be CM if A is isogenous to a product of simple CM Abelian varieties. If $G = GS_p(2g, \mathbb{Q})$ (as in Example 3.14) and $x \in \mathbb{H}_g$. Then $x(\mathbb{S})$ is contained in the real points $T(\mathbb{R})^+$ of a $\mathbb{Q}$-torus $T_{\mathbb{Q}}$ of $G_{\mathbb{Q}}$ if and only if the image of x in $S = \mathcal{A}_{g,N}$ corresponds to a CM Abelian variety. A similar description in terms of endomorphism of Hodge structure exists for a general Shimura variety. Therefore a special point is often called a CM point.

With these definitions one can check that special points are dense in the Zariski topology in any special variety (they are in fact dense in the analytic topology). The existence of one CM point is given by the study of the space $\mathcal{T}_G$ of maximal tori of $G_{\mathbb{Q}}$. Suppose that G is semi-simple. It can be shown that $\mathcal{T}_G$ is a rational variety and that the locus of compact tori is open in the usual topology. Any $\mathbb{Q}$-rational point of $\mathcal{T}_G$ which is in this open set will define a CM point. Note that if $x \in S$ is special then for all $g \in G(\mathbb{Q})$, $T_g.x$ is a finite union of special points and the union of the $T_g.x$, for $g \in G(\mathbb{Q})$ is dense in the analytic topology of S.

A component of the intersection of special subvarieties is special (this is not clear from the point of view of Hermitian locally symmetric spaces, but from the moduli point of view the intersection is interpreted as the locus of points with the additional structures of all the subvarieties we are intersecting).

A component of the image by a Hecke operator of a special variety is a special variety.

EXAMPLE 3.17. The special subvarieties of $S = \mathrm{SL}(2, \mathbb{Z})\backslash\mathbb{H}$ are S and the CM points corresponding to the CM elliptic curves studied in Section 2.3.

EXAMPLE 3.18. If S is a Shimura variety any Hecke correspondence T_q is a special subvariety of the Shimura variety $S \times S$.

EXAMPLE 3.19. The j-function induces an isomorphism $S = \mathrm{SL}(2, \mathbb{Z})\backslash \mathbb{H} \simeq \mathbb{C}$. The special subvarieties of $\mathbb{C} \times \mathbb{C}$ are

(i) $S \simeq \mathbb{C}$.

(ii) Couples of CM points .

(iii) Curves of the form $\{x\} \times \mathbb{C}$ (or $\mathbb{C} \times x$) for some CM point x.

(iv) Modular correspondences $Y_0(N) = \Gamma_0(N)\backslash\mathbb{H}$ associated to $q_N = \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}(2, \mathbb{Q})$: $Y_0(N)$ is the image in $\mathbb{C} \times \mathbb{C}$ of the map $(j(\tau), j(N\tau))$. $Y_0(N)$ is the (coarse) moduli space for triples $(E, E', \alpha : E \to E')$ where α is a cyclic isogeny of degree N.

EXAMPLE 3.20. For each totally real number field F of degree g the associated Hilbert modular varieties are special subvarieties of Siegel modular varieties. This is clear from the "moduli interpretation" and the translation in terms of Shimura data is for example explained in (Van der Geer, 1988, Chap. 9.2). Any component of a Hecke translate of an Hilbert modular variety is again an Hilbert modular variety associated to an order A of O_F.

The Conjecture 3.2 in this case was formulated by André for a curve in a general Shimura variety and by Oort for subvarieties of arbitrary dimension in $\mathcal{A}_g$.

CONJECTURE 3.21 (André–Oort).

(a) *An irreducible component of the Zariski closure of a set of* CM *points is a special subvariety.*

(b) *Let Y be an algebraic subvariety of X. There exists special subvarieties $\{Z_1, \ldots, Z_r\}$ with $Z_i \subset Y$ such that if $Z \subset Y$ is a special subvariety then*

$$Z \subset \bigcup_{i=1}^{r} Z_i.$$

A deep fact of the theory of Shimura varieties is that any Shimura variety is "canonically" defined over a number field. In particular CM points are defined over $\overline{\mathbb{Q}}$. One proof of this fact is given by Faltings (Faltings, 1984a) using a rigidity argument. Another proof is an important achievement of the work of several mathematicians. Let's just mention among them Shimura, Deligne, Borovoi, Milne, Shih (Milne and Shih, 1982). The fundamental fact which is not given by Faltings's approach is the knowledge of the field of definition (the reflex field) of the Shimura variety which can be computed in terms of the Shimura datum (G, X). Note that if you are allowed to use the adeles (as in Deligne's approach) then the Shimura variety

$$\mathrm{Sh}_K(G, X) = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/K$$

(where K is a compact open subgroup of $G(\mathbb{A}_f)$), is canonically defined over the reflex field. With our definition, a Shimura variety is a connected component of $\mathrm{Sh}_K(G, X)$ and is defined over an Abelian extension of the reflex field.

Moreover a special subvarieties Z of a Shimura variety is a Hermitian locally symmetric space, therefore Z is endowed with a canonical probability measure μ_Z such that $\mathrm{Supp}(\mu_Z) = Z$.

The set $\mathcal{S}(S)$ of special subvarieties is therefore "strongly admissible" (with the terminology of Section 3.1). The general equidistribution conjecture can then be stated:

CONJECTURE 3.22. *The subset $\mathcal{Q}(S)$ of $\mathcal{P}(S)$ is compact. If Δ_{Z_n} is a sequence of measure in $\mathcal{Q}(S)$ weakly converging to μ_Z then for all $n \gg 0$,* $\mathrm{Supp}(\mu_{Z_n}) \subset \mathrm{Supp}(\mu_Z)$.

The evidence for this conjecture is limited. The only known case is the theorem of Duke explained in Section 2.2.3: for $S = \mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H}$, the Galois orbits of CM points are equidistributed. The case of a Shimura curve associated to a quaternion algebra over $\mathbb{Q}$ is almost known by some work of Zhang (Zhang, 2005). General results for equidistribution of the sequence of probability measure associated to *connected* special subvarieties are obtained using ergodic theory and will be described in section 4.3.

The ideas from the Arakelov theory (successful in the case of Abelian varieties as described in Section 3.2) are not applicable here because for any reasonable theory of height, the height of CM points is unbounded (see (Colmez, 1998) for the case of CM elliptic curves).

3.3.1. *Equidistribution of "Toric orbits" of* CM *points*

No general strategy for general Shimura varieties is known for the Conjecture 3.22. The ideas behind Duke's proof of the case of CM elliptic curves can be extended to the case of Hilbert modular varieties (Clozel and Ullmo, 2005a; Cohen, 2005) and (Venkatesh, 2005) or more generally for Shimura varieties associated to quaternion algebra over a totally real number field (Zhang, 2005). This extension unfortunately doesn't lead to the equidistribution of Galois orbits of CM points (and therefore to the André–Oort conjecture) but leads to the equidistribution of a bigger set: the *Toric orbits of* CM *points.* Some attempts to obtain in some cases equidistribution of Galois orbits of CM points are given in (Venkatesh, 2005; Zhang, 2005). See Section 4.2.2.

4. Equidistribution of Special Subvarieties

4.1. THE CASE OF ABELIAN VARIETIES

Let $A = \Gamma\backslash\mathbb{C}^n$ be a complex Abelian variety. An Abelian subvariety B of A is canonically endowed with a probability measure μ_B such that $\mathrm{Supp}(\mu_B) = B$.

Let $\mathcal{P}(A)$ be the set of Borel probability measures on A and

$$Q(A) = \{\mu_B, \text{ with } B \text{ an Abelian subvariety of } A\}.$$

PROPOSITION 4.1. *The set $Q(A)$ is compact and if μ_{B_n} is a sequence of $Q(A)$ weakly converging to μ_B then for all $n \gg 0$, B_n is an Abelian subvariety of B.*

A sequence B_n of Abelian subvarieties of A is said to be strict (relatively to A) if for all proper Abelian subvariety B of A the set $\{n \in \mathbb{N} \mid B_n \subset B\}$ is finite. For formal reasons the proposition is equivalent to

PROPOSITION 4.2. *Let B_n be a strict sequence of Abelian subvarieties of A then μ_{B_n} weakly converges to μ_A.*

Let's recall why the two propositions are equivalent: let B_n be a sequence of Abelian subvarieties. Let $\mathcal{E}$ be the set of Abelian subvarieties of A containing infinitely many B_n's. Let A' be a minimal element of $\mathcal{E}$. Replacing B_n by a subsequence we can suppose that B_n is a strict sequence of A' and the Proposition 4.2 implies that μ_{B_n} is weakly converging to $\mu_{A'}$. The implication "Proposition 4.1 implies Proposition 4.2" is simpler (and left as an exercise).

The proof will use only classical Fourier theory. The first step consists in "forgetting the complex structures":

4.1.1. *The flat case*

In this part we write $G = \mathbb{Q}^n$, $X = \mathbb{Z}^n\backslash\mathbb{R}^n$ and $\pi\colon \mathbb{R}^n \rightarrow$ the canonical morphism. So X is just a C^∞-variety but we define a set $\mathcal{S}(X)$ of special subvarieties by

$$\mathcal{S}(X) = \{Z = \pi(H \otimes_{\mathbb{Q}} \mathbb{R}) \text{ with } H \text{ a } \mathbb{Q}\text{-vector subspace of } G\}.$$

Every $Z = \pi(H \otimes_{\mathbb{Q}} \mathbb{R})$ is canonically endowed with a probability measure coming from the Lebesgue measure on $H \otimes_{\mathbb{Q}} \mathbb{R}$.

In this situation we can formulate the analogue of Propositions 4.1 and 4.2. As in the previous part the 2 statements are equivalent

The purpose of this part is to prove the analogue of Proposition 4.2:

PROPOSITION 4.3. *Let Z_n be a strict sequence* (*relatively to the set of special subvarieties*) *of special subvarieties of X then μ_{Z_n} weakly converges to μ_X.*

For $x \in \mathbb{R}^n$ we write $\bar{x}$ for the class of x in $\mathbb{Z}^n\backslash\mathbb{R}^n$. The set X^* of complex character of X is in bijection with $\mathbb{Z}^n$: for all $\underline{\mathrm{k}} = (k_1, \ldots, k_n) \in \mathbb{Z}^n$ the associated character $\chi_{\underline{\mathrm{k}}}$ of X is defined by

$$\chi_{\underline{\mathrm{k}}}(\bar{x}_1, \ldots, \bar{x}_n) = \exp\left(2i\pi \sum_{j=1}^{n} k_j x_j\right).$$

Any character of X is obtained in this way.

If $\chi = \chi_{k_1,\ldots,k_n}$ for some $(k_1, \ldots, k_n) \in \mathbb{Z}^n - \{(0, \ldots, 0)\}$, we write

$$H_\chi = H_{k_1,\ldots,k_n}$$

the $\mathbb{Q}$-hyperplane of G defined by

$$\sum_{j=1}^{n} k_j x_j = 0.$$

Then $H_\chi = H_{\chi'}$ with $\chi = \chi_{k_1,\dots,k_n}$ and $\chi' = \chi_{k'_1,\dots,k'_n}$ if and only if there exists $\alpha \in \mathbb{Q}$ such that $k'_i = \alpha k_i$ for all i.

Let

$$S_\chi = S_{k_1,\dots,k_n} = \pi(H_{k_1,\dots,k_n} \otimes \mathbb{R})$$

be the associated maximal special subvariety of X. All the maximal special subvarieties are obtained in this way. We define also

$$\widetilde{S_\chi} = H_{k_1,\dots,k_n} \otimes \mathbb{R}.$$

LEMMA 4.4. *Let S be a special subvariety of X and $\chi = \chi_{k_1,\dots,k_n}$ a non-trivial character of X. The restriction χ_S of χ on S is a character and $\chi_S = 1$ if and only if $S \subset S_\chi$.*

The character χ is trivial on S_χ, therefore if $S \subset S_\chi$ then $\chi_S = 1$. If $\chi_S = 1$, $S = \pi(\widetilde{S})$ for a $\mathbb{R}$-subvector space of $G(\mathbb{R})$ and $x = (x_1, \dots, x_n) \in \widetilde{S}$ then for all $t \in \mathbb{R}$, $t\sum_{i=1}^n k_i x_i \in \mathbb{Z}$. Therefore $\sum_{i=1}^n k_i x_i = 0$ and $\pi(x) \in S_\chi$ and $S \subset S_\chi$.

We can now give a proof of Proposition 4.3. Let T_n be a strict sequence of special subvarieties of X and μ_n the associated sequence of probability measure. Using Weyl's criterion, we must show that for all non-trivial character of X

$$\lim_{n\to\infty} \int_{T_n} \chi \, d\mu_n = \lim_{n\to\infty} \int_{T_n} \chi_{T_n} \, d\mu_n = 0.$$

Fix a character χ. As T_n is a strict sequence, for all n big enough T_n is not contained in S_χ. By Lemma 4.4 χ_{T_n} is a non-trivial character of T_n and

$$\int_{T_n} \chi_{T_n} \, d\mu_n = 0.$$

EXERCISE 4.5. Prove the following analogue of Proposition 4.1: let S_n be a sequence of special subvarieties of $X = \mathbb{Z}^n \backslash \mathbb{R}^n$. Then there exists a special subvariety S and a subsequence S_{n_k} such that μ_{n_k} converges weakly to μ_S. Moreover for all $k \gg 0$, $S_{n_k} \subset S$.

REMARK 4.6. We can replace $\mathbb{Z}^n$ by an arbitrary lattice Γ of $\mathbb{R}^n$ in the previous statements. There exists a linear automorphism u_Γ of $\mathbb{R}^n$ such that $u_\Gamma(\mathbb{Z}^n) = \Gamma$. Such an automorphism induces an isomorphism of C^∞ varieties;

$$\overline{u_\Gamma}\colon \mathbb{Z}^n \backslash \mathbb{R}^n \to \Gamma \backslash \mathbb{R}^n.$$

The special subvarieties of $\Gamma \backslash \mathbb{R}^n$ are just the images by $\overline{u_\Gamma}$ of the special subvarieties of $\mathbb{Z}^n \backslash \mathbb{R}^n$.

We can now give a proof of Proposition 4.2. Let $A = \Gamma\backslash\mathbb{C}^n$ an Abelian variety. Identifying $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ and applying the previous remark we have a notion of special subvarieties of A. (*Warning*: this notion is not the usual one for Abelian varieties). With this notion, Abelian subvarieties of A are special subvarieties but there exists *much more special subvarieties* corresponding to non-complex subtori. (For example if A is a simple Abelian variety, there exists no sub-Abelian varieties).

Let A_n be a sequence of Abelian subvarieties of A and μ_n the associated sequence of probability measures. Then using Exercise 4.5, there exists a special subvariety S of A and a subsequence μ_{n_k} weakly converging to μ_S. For all k big enough $A_{n_k} \subset S$. We therefore just need to prove that S is an Abelian subvariety. The group generated by the A_{n_k} for $k \gg 0$ is generated by finitely many Abelian subvarieties and is therefore an Abelian subvariety B. So $B \subset S$, but the supports of the μ_{n_k} for $k \gg 0$ are contained in B which is a closed subvariety. As $\mu_{n_k} \to \mu_S$, we find that $S = \mathrm{supp}(\mu_S) \subset B$ and therefore that $S = B$ is an Abelian subvariety of A.

4.2. EQUIDISTRIBUTION OF ALGEBRAIC MEASURES

Let $G_{\mathbb{Q}}$ be a connected algebraic group over $\mathbb{Q}$ and $X^*(G_{\mathbb{Q}})$ be the set of rational characters of $(G_{\mathbb{Q}})$. We say that $G_{\mathbb{Q}}$ is of type $\mathcal{F}$ if $X^*(G_{\mathbb{Q}}) = \{1\}$. Fix a $\mathbb{Z}$-structure $G_{\mathbb{Z}}$ on $G_{\mathbb{Q}}$ (for example the Zariski closure of $G_{\mathbb{Q}}$ in $\mathrm{GL}(n, \mathbb{Z})$ for a faithful representation of G in $\mathrm{GL}(n, \mathbb{Q})$). A subgroup $\Gamma \subset G(\mathbb{Q})^+$ is said to be an "arithmetic lattice" if Γ is commensurable with $G_{\mathbb{Z}}(\mathbb{Z})$ (this doesn't depends of the choice of the $\mathbb{Z}$-structure on $G_{\mathbb{Q}}$).

Let G denote the real Lie group $G = G(\mathbb{R})^+$ and let μ_G be the G-invariant measure on $X^+ = \Gamma\backslash G$. Then the μ_G-volume of X^+ is finite, hence the name lattice, see (Platonov and Rapinchuk, 1994, Theorem 4.13, p. 213).

If $H_{\mathbb{Q}} \subset G_{\mathbb{Q}}$ is a connected $\mathbb{Q}$-algebraic subgroup of type $\mathcal{F}$, then $\Gamma_H = \Gamma \cap H(\mathbb{R})^+$ is an arithmetic lattice of $H = H_{\mathbb{Q}}(\mathbb{R})^+$ and

$$X_H^+ = \Gamma\backslash\Gamma H(\mathbb{R})^+ \simeq \Gamma_H\backslash H(\mathbb{R})^+$$

is a closed subset of X^+ endowed with a canonical H-invariant probability measure μ_H. Such a subset is said to be special in this section. A probability measure on X^+ is said to be "algebraic" (or homogeneous) if it is obtained in this way.

Let $\mathcal{P}(X)^+$ be the set of probability measures on X^+ endowed with the weak star topology. Let $Q(X^+)$ be the subset of $\mathcal{P}(X^+)$ consisting of the algebraic measures. There is a natural definition of strict sequence of $\mathbb{Q}$-subgroups of $G_{\mathbb{Q}}$: such a sequence $H_{n,\mathbb{Q}}$ is said to be strict if for all proper $\mathbb{Q}$-subgroup $H_{\mathbb{Q}}$ the set

$$\{n \in \mathbb{N}, H_{n,\mathbb{Q}} \subset H_{\mathbb{Q}}\}$$

is finite.

We'll give some general examples of such strict sequence $H_{n,\mathbb{Q}}$ verifying the following equidistribution property:

($\mathcal{E}$) The associated sequence of probability measures $\mu_n = \mu_{H_n}$ weakly converges to μ_G.

The following example shows that the property ($\mathcal{E}$) is not always verified:

EXAMPLE 4.7. Let $G_{\mathbb{Q}}$ be the group SU(2). Therefore $G = G(\mathbb{R})^+$ is compact and any (arithmetic) lattice Γ is finite. Therefore we may assume that $\Gamma = \{1\}$ and $X^+ = G$. Fix a $\mathbb{Q}$-torus $T_{0,\mathbb{Q}}$ of $G_{\mathbb{Q}}$, and a sequence $g_n \in G(\mathbb{Q})$ converging to $g \in G(\mathbb{R})$. Note that the canonical measure μ_{T_0} is just the normalized Haar measure on $T_0(\mathbb{R})^+$ in this situation. Suppose that the sequence $T_{n,\mathbb{Q}} = g_n T_{0,\mathbb{Q}} g_n^{-1}$ is strict (as an exercise prove that this is possible). Then μ_{T_n} is weakly convergent to the normalized Haar measure on $gT^0(\mathbb{R})^+ g^{-1} \neq \mu_G$.

We proposed in (Clozel and Ullmo, 2005a, with Laurent Clozel) the following conjecture for which we don't know any counter-example. The formulation uses the adeles. One of the assignments of the organizers was to "avoid the adeles like the plague". So if you are afraid of contamination you should avoid this part.

Let $\mathbb{A}$ be the ring of adeles of $\mathbb{Q}$ and $\mathbb{A}_f$ the ring of finite adeles. Let $G_{\mathbb{Q}}$ be an algebraic group of type $\mathcal{F}$. A congruence subgroup of $G(\mathbb{Q})$ is an arithmetic lattice Γ of $G(\mathbb{R})^+$ of the form

$$\Gamma = G(\mathbb{Q})^+ \cap K$$

for an open compact subgroup K of $G(\mathbb{A}_f)$. Let Γ be a congruence subgroup of $G(\mathbb{Q})$ and $X = \Gamma \backslash G(\mathbb{R})^+$. Then X is a component of

$$S(G, K) = G(\mathbb{Q})^+ \backslash G(\mathbb{R})^+ \times G(\mathbb{A}_f)/K$$

and the components of $S(G, K)$ are indexed by the finite set $G(\mathbb{Q})^+ \backslash G(\mathbb{A}_f)/K$ (which should be thought of as a class-group).

If $H_{\mathbb{Q}}$ is a $\mathbb{Q}$-subgroup of type $\mathcal{F}$, then $K_H = H(\mathbb{A}_f) \cap H$ is an open compact subgroup of $H(\mathbb{A}_f)$. Every irreducible component of

$$S(H, K) = S(H, K_H) = H(\mathbb{Q})^+ \backslash H(\mathbb{R})^+ \times H(\mathbb{A}_f)/K_H$$

is endowed with a canonical probability measure. Let $\Theta_{X,H}$ be the set of components of $S(H, K_H)$ which are contained in X and let $h_a = |\Theta_{X,H}|$. The adelic probability measure $\mu_{a,H}$ associated to H is by definition:

$$\mu_{a,H} = \frac{1}{h_a} \sum_{\gamma \in \Theta_{X,H}} \mu_\gamma$$

where μ_γ is the canonical probability measure on the component Z_γ of $S(H, K_H)$ indexed by $\gamma \in \Theta_{X,H}$. The adelic equidistribution conjecture is then:

CONJECTURE 4.8. *Let H_n be a strict sequence of $\mathbb{Q}$-subgroups of $G_{\mathbb{Q}}$ of type $\mathcal{F}$. Then the associated sequence of measures μ_{a,H_n} weakly converges to μ_G.*

In the Example 4.7 it can be shown that the cardinality of the class group $T_n(\mathbb{Q})^+\backslash T_n(\mathbb{A}_f)/K \cap T_n(\mathbb{A}_f)$ tends to ∞ as n tends to ∞.

4.2.1. *Ergodic theory and property $\mathcal{E}$*

In this section we explain a general situation where the property $\mathcal{E}$ is verified. We start by the following definition:

DEFINITION 4.9. A connected linear algebraic group over $\mathbb{Q}$ is said to be of type $\mathcal{H}$ if its solvable radical is unipotent and if $H_s = H/R_u(H)$ is an almost direct product of $\mathbb{Q}$-simple groups H_i such that $H_i(\mathbb{R})$ is not compact.

THEOREM 4.10. *Let $G_{\mathbb{Q}}$ be a semi-simple group of type $\mathcal{H}$, and $H_n \subset G_{\mathbb{Q}}$ be a strict sequence of subgroups of type $\mathcal{H}$. Then the property ($\mathcal{E}$) is verified for the associated sequence of measure $\mu_n = \mu_{H_n}$.*

Let's give some ideas of the proof of such a result. Let $G_{\mathbb{Q}}$ be as in the statement of the theorem and $\Gamma \in G(\mathbb{Q})^+$ be an arithmetic lattice and $X = \Gamma\backslash G(\mathbb{R})^+$.

DEFINITION 4.11. Let $F \subset G(\mathbb{R})^+$ be a connected closed Lie subgroup. We say that F is of type $\mathcal{K}$ if

(i) $F \cap \Gamma$ is a lattice in F. Therefore $F \cap \Gamma\backslash F$ is a closed subset of $\Gamma\backslash G(\mathbb{R})^+$. Let μ_F be the associated F-invariant probability measure.

(ii) The subgroup $L(F)$ generated by the unipotent one parameter subgroup of F acts *ergodicaly* on $F \cap \Gamma\backslash F$ with respect to μ_F. By definition this means that any $L(F)$-invariant measurable subset of $F \cap \Gamma\backslash F$ is of μ_F-measure 0 or 1.

The relation between the class $\mathcal{K}$ and the class $\mathcal{H}$ is given by the following lemma (see (Clozel and Ullmo, 2005a) for a proof).

LEMMA 4.12.

(a) *If $H_{\mathbb{Q}}$ is an algebraic $\mathbb{Q}$-subgroup of type $\mathcal{H}$ then $H(\mathbb{R})^+$ is a Lie subgroup of type $\mathcal{K}$.*

(b) *If F is a closed Lie subgroup of type $\mathcal{K}$, then there exists an algebraic $\mathbb{Q}$-subgroup $F_{\mathbb{Q}}$ of type $\mathcal{H}$ such that $F = F(\mathbb{R})^+$.*

The algebraic group $F_{\mathbb{Q}}$ associated to F in this last statement is the "Mumford-Tate group of F": $F_{\mathbb{Q}}$ is the smallest $\mathbb{Q}$-subgroup $H_{\mathbb{Q}}$ of $G_{\mathbb{Q}}$ such that $F \subset H(\mathbb{R})^+$.

A deep result of Ratner (Ratner, 1991a; Ratner, 1991b) (conjectured by Raghunathan) implies that if L is a closed Lie subgroup of $G(\mathbb{R})^+$ generated by one parameter unipotent subgroups then the Mumford-Tate group $F_{\mathbb{Q}}$ of L is of type $\mathcal{H}$ and the closure $\overline{\Gamma\backslash\Gamma L}$ in the analytic topology of $\Gamma\backslash\Gamma L$ is $\Gamma\backslash\Gamma F(\mathbb{R})^+ \simeq F(\mathbb{R})^+ \cap \Gamma\backslash F(\mathbb{R})^+$.

Let $\mathcal{P}(X)$ be the set of Borel probability measure on X and $Q(X)$ be the subset of $\mathcal{P}(X)$ defined as

$$Q(X) = \{\mu_F, F \in \mathcal{K}\}.$$

As a consequence of the previously discussed work of Ratner, Mozes–Shah (Mozes and Shah, 1995) proved the following (deep) analogue of Proposition 4.1:

THEOREM 4.13 (Mozes-Shah). *The set $Q(X)$ is compact in the weak star topology. If μ_n is a sequence of $Q(X)$ weakly converging to μ, then $\mu \in Q(X)$ and for all n big enough* $\mathrm{Supp}(\mu_n) \subset \mathrm{Supp}(\mu)$.

The proof of Theorem 4.10 is now straightforward. Let $H_{n,\mathbb{Q}}$ be a sequence of algebraic subgroups of $G_{\mathbb{Q}}$ of type $\mathcal{H}$ and $\mu_n \in Q(X)$ be the associated sequence. If μ_α is a subsequence converging to μ, then $\mu = \mu_H$ for a closed connected Lie subgroup of type $\mathcal{K}$. Then $H = H_{\mathbb{Q}}(\mathbb{R})^+$ for an algebraic $\mathbb{Q}$-subgroup of type $\mathcal{H}$. For all $\alpha \gg 0$, $\mathrm{Supp}(\mu_\alpha) \subset \mathrm{supp}(\mu)$, therefore $\mathrm{Lie}(H_\alpha(\mathbb{R})) \subset \mathrm{Lie}(H(\mathbb{R}))$. Hence $H_\alpha(\mathbb{R})^+ \subset H(\mathbb{R})^+$ and by the definition of the Mumford–Tate group $H_{\alpha,\mathbb{Q}} \subset H_{\mathbb{Q}}$. As the sequence $H_{n,\mathbb{Q}}$ is strict $H_{\mathbb{Q}} = G_{\mathbb{Q}}$ and $\mu = \mu_G$.

4.2.2. *Adelic equidistribution for* PGL(2, F)

In view of the last section the property $\mathcal{E}$ may fail for sequences $H_{n,\mathbb{Q}}$ of reductive non-semi-simple algebraic subgroups of $G_{\mathbb{Q}}$. The following statement (Clozel and Ullmo, 2005a, Theorem 7.1) is an important case where the property $\mathcal{E}$ may fail, but the adelic equidistribution conjecture 4.8 holds.

THEOREM 4.14. *Let F be a number field and $G_{\mathbb{Q}} = \mathrm{Res}_{F/\mathbb{Q}}\,\mathrm{PGL}(2, F)$. Let O_F be the ring of integers of F and d_n be a sequence of square free elements of O_F. Then*

$$T'_{n,\mathbb{Q}} = \left\{\begin{pmatrix} a & b \\ d_n b & a \end{pmatrix}, a^2 - d_n b^2 \neq 0\right\}$$

is a torus of $\mathrm{Res}_{F/\mathbb{Q}}\,\mathrm{GL}(2,F)$*. Let* $T_{n,\mathbb{Q}}$ *be the image of* $T'_{n,\mathbb{Q}}$ *in* $G_{\mathbb{Q}}$*. Let* Γ *be a congruence subgroup of* $G(\mathbb{Q})^+$ *and* $X = \Gamma\backslash G(\mathbb{R})^+$*. If the norm* $N_{F/\mathbb{Q}}(d_n)$ *of* d_n *tends to* ∞*, then the associated adelic measure* $\mu_{a,n} = \mu_{a,T_n}$ *weakly converges to* μ_G*.*

We only give a description of the proof which is an extension of Duke's method for the equidistribution of CM points on $\mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H}$ discussed in Section 2.2.3.

Let f be a parabolic form on X, π the associated automorphic representation. For $d \in O_F$ with d square free, we denote by Π_d the base change of π to $E_d = F[\sqrt{d}]$. Using a formula of Waldspurger (Waldspurger, 1985, Proposition 7, p. 222), we obtain a relation between $\mu_{a,d}(f)$ and $L(\Pi_d, \frac{1}{2})$ and we show that for all $\varepsilon > 0$

$$|\mu_{a,d}(f)| \ll |N_{F/\mathbb{Q}}(d)|^{(-1/4)+(\theta/2)+\varepsilon}, \tag{21}$$

where θ denotes the "Selberg constant" ($0 \leq \theta < \frac{1}{2}$) measuring the lack of validity of the Selberg conjecture (predicting $\theta = 0$). The Lindelöf hypothesis combined with the Selberg conjecture would give:

$$|\mu_{a,d}(f)| \ll |N_{F/\mathbb{Q}}(d)|^{-1/2+\varepsilon}.$$

The same kind of results is obtained for Eisenstein series $E_\chi(g,s)$ associated with a character χ of $F^*\backslash\mathbb{A}_f^*$ using a result of Wielonski (Wielonsky, 1985): $\mu_{a,d}(E_\chi(g, \frac{1}{2} + i\sigma))$ is related to the special value of a L-function "à la Tate" $L(\chi N_{E_d/F}, \frac{1}{2})$.

There exists $A > 0$ such that for all $\varepsilon > 0$ and all $\sigma \in \mathbb{R}$:

$$\left|\mu_{a,d}\Big(E_\chi(g, \tfrac{1}{2} + i\sigma)\Big)\right| \ll |N_{F/\mathbb{Q}}(d)|^{(-1/4)+\varepsilon}|\sigma|^A. \tag{22}$$

The Lindelöf hypothesis for $L(\chi \circ N_{E/F}, \frac{1}{2})$ would give

$$\left|\mu_{a,d}\Big(E_\chi(g, \tfrac{1}{2} + i\sigma)\Big)\right| \ll |N_{F/\mathbb{Q}}(d)|^{(-1/2)+\varepsilon}|\sigma|^A.$$

Note that we don't need a subconvexity result in the proof of the theorem. The method of the proof leads to a conditional statement for the analogue statement on the symmetric space. In this case we lose a power of $N_{F/\mathbb{Q}}(d)$ the bounds for $|\mu_{a,d}(f)|$ and $\left|\mu_{a,d}\Big(E_\chi(g, \frac{1}{2} + i\sigma)\Big)\right|$ and we need a subconvexity bound as in Duke's theorem.

For example if F is totally real, and T_n is a sequence of tori such that $T(\mathbb{R})$ is compact this is the problem of equidistribution of toric orbits of CM points on a Hilbert modular variety discussed briefly in Section 3.3.1.

Venkatesh (Venkatesh, 2005) has a method which leads to unconditional results in this case.

Note that from the harmonical analysis point of view the situation is much harder than in the case of Duke (say $F = \mathbb{Q}$). Using equations (21) and (22) you get a L^2-convergence and you want to deduce from this a pointwise convergence (see Section 2.3 for a detailed similar example). In Duke's case the continuous part of $L^2(\mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H}, d\mu_0)$ is obtained using one Eisenstein series. For a general number field you need to consider an infinite set of Eisenstein series (essentially parametrised by the units of O_F). You therefore need to understand the dependence in χ of the bounds given in the equation (22).

4.3. EQUIDISTRIBUTION OF SPECIAL SUBVARIETIES OF SHIMURA VARIETIES

The references for this part are (Clozel and Ullmo, 2005b; Ullmo, 2006).

Let S be a connected Shimura variety as defined in Section 3.3. Let Z_n be a sequence of special subvarieties of S. You can't expect in general that the associated sequence of probability measure $\mu_n = \mu_{Z_n}$ weakly converges. For example if x_n is a sequence of CM points then μ_n is just the Dirac measure supported at x_n, such a sequence can converge to δ_x for a non-CM point or x_n may tend to ∞. Even for positive dimensional special subvarieties the same problem may occur. Start with a special subvariety $Z \times Z'$ of S for two special varieties Z and Z'. If x_n is a CM point of Z' and $Z_n = Z \times \{x_n\}$ there is no hope of proving the week convergence of μ_n. A special subvariety Z of S is said to be NF (non-factor) if Z is not of the form $Z_1 \times \{x\}$ with Z_1 special and x a CM point. The following theorem is obtained in (Ullmo, 2006) and is a generalization of the main result of (Clozel and Ullmo, 2005b) obtained with L. Clozel.

THEOREM 4.15. *Let $\mathcal{P}(S)$ be the set of Borel probability measure on S. Let $\mathcal{Q}(S) = \{\mu_Z \mid Z\mathrm{NF}\}$. Then $\mathcal{Q}(S)$ is compact and if μ_n is a sequence of $\mathcal{Q}(S)$ converging to μ, then $\mu = \mu_Z \in \mathcal{Q}(S)$ and for all $n \gg 0$, $Z_n \subset Z$.*

As usual, we get the following result in the direction of the André–Oort conjecture .

THEOREM 4.16. *Let $Y \subset S$ be a subvariety of a Shimura variety S. Then there exists a finite set $\{Z_1, \ldots, Z_r\}$ of* NF *special subvarieties of Y such that if Z is a special* NF *subvariety of Y then*

$$Z \subset \bigcup_{i=1}^{r} Z_i.$$

Let's recall how Theorems 4.15 implies 4.16. Let Z_n be a sequence of distinct NF special subvarieties of Y which are maximal among NF special subvarieties of Y. By Theorem 4.15 we can suppose that the associated sequence of probability measure μ_n converges weakly to μ_Z for a NF special subvariety of S. But as Y is closed $\mathrm{supp}(\mu_Z) = Z \subset Y$. For n big enough we know that $Z_n \subset Z \subset Y$.

Let's give a sketch of the proof of the Theorem 4.15 (the reader should compare with the proof of the Proposition 4.1). There exists a semi-simple group $G_{\mathbb{Q}}$ such that the associated symmetric space $\mathcal{D} = G(\mathbb{R})^+/K_\infty$ is Hermitian and an arithmetic lattice $\Gamma \in G(\mathbb{Q})^+$ such that $S = \Gamma\backslash\mathcal{D}$.

Let $\Omega = \Gamma\backslash G(\mathbb{R})^+$. We defined in Section 4.1.1, a class $\mathcal{H}$ of algebraic $\mathbb{Q}$-subgroups of $G_{\mathbb{Q}}$ and a compact subset $Q(\Omega)$ of the set $\mathcal{P}(\Omega)$ of Borel probability measure on Ω. A special subvariety Z of S is associated to a $\mathbb{Q}$-subgroup $H_{\mathbb{Q}}$ such that $H_{\mathbb{Q}}^{\mathrm{der}}$ is of type $\mathcal{H}$.

Let Z_n be a sequence of special subvarieties and $H_{n,\mathbb{Q}}$ be the associated sequence. Suppose for simplicity that $H_{n\mathbb{Q}}$ is semi-simple (this is the case considered in (Clozel and Ullmo, 2005b)), then $H_{n\mathbb{Q}}$ is of type $\mathcal{H}$ Using the results of Mozes and Shah (Theorem 4.13), we may assume that the associated sequence μ_n of $Q(\Omega)$ weakly converges to a measure $\mu_H \in Q(\Omega)$ for an algebraic $\mathbb{Q}$-subgroup $H_{\mathbb{Q}}$ of $G_{\mathbb{Q}}$ of type $\mathcal{H}$. For all $n \gg 0$ we know moreover that $H_{n,\mathbb{Q}} \subset H_{\mathbb{Q}}$. We then show that $H_{\mathbb{Q}}$ is related to Shimura varieties: $H_{\mathbb{Q}}$ should be reductive (and in fact semi-simple in view of the Definition 4.9 of type $\mathcal{H}$) and the symmetric space associated to $H_{\mathbb{R}}$ should be of Hermitian type. (The formalism of Deligne (Deligne, 1971; Deligne, 1979) is used in this part).

You need then to pass from this result on $\Omega = \Gamma\backslash G(\mathbb{R})^+$ to a result on the Shimura variety $S = \Gamma\backslash G(\mathbb{R})^+/K_\infty$. The main difficulty is the following: for each point $x \in \mathcal{D}$ we have an associated maximal compact subgroup K_x of $G(\mathbb{R})^+$ and a morphism $\pi_x\colon \Omega \to S$. Let Z be a special subvariety of S with associated canonical probability measure μ_Z. Let $H_{\mathbb{Q}}$ an associated $\mathbb{Q}$-subgroup and μ_H as previously. To understand the relation between μ_Z and μ_H, you must fix a maximal compact subgroup K_x of $G(\mathbb{R})^+$ such that $H(\mathbb{R})^+ \cap K_x$ is a maximal compact subgroup of $H(\mathbb{R})^+$. Then $\mu_Z = \pi_{x\star}\mu_H$. If you could fix a x such that $K_x \cap H_n(\mathbb{R})^+$ is a maximal compact subset of $H_n(\mathbb{R})^+$ for all $n \in \mathbb{N}$ then the result on Ω would directly imply the Theorem 4.16. To overcome this difficulty (which is not serious if Γ is a cocompact lattice of $G(\mathbb{R})^+$—i.e. if $G_{\mathbb{Q}}$ is $\mathbb{Q}$-anisotropic), we use some results of Dani and Margulis on the behavior of unipotent flows (Dani and Margulis, 1991).

References

Abbes, A. (1997) Hauteurs et discrétude (d'après L. Szpiro, E. Ullmo et S. Zhang), In *Séminaire Bourbaki, vol. 1996/97*, pp. 141–166.

Autissier, P. (2004) Equidistribution des sous-variétés de petite hauteur des variétés abéliennes, Technical report, IRMA 04-22.

Baker, M. and Ih, S. (2004) Equidistribution of small subvarieties of an Abelian variety, *New York J. Math.* **10**, 279–289.

Bilu, Y. (1997) Limit distribution of small points on algebraic tori, *Duke Math. J.* **89**, 465–476.

Borel, A. (1969) *Introduction aux groupes arithmétiques*, Vol. 1341 of *Actualités Sci. Indust.*, Paris, Hermann.

Bump, D., Duke, W., Hoffstein, J., and Iwaniec, H. (1992) An estimate for the Hecke eigenvalues of Maass forms, *IMRN* **4**, 75–81.

Chambert-Loir, A. (2000) Points de petite hauteur sur les variétés semi-abéliennes, *Ann. Sci. École Norm. Sup. (4)* **33**, 789–821.

Clozel, L., Oh, H., and Ullmo, E. (2001) Hecke operators and equidistribution of Hecke points, *Invent. Math.* **144**, 327–351.

Clozel, L. and Ullmo, E. (2001) Equidistribution des points de Hecke, In Hida, Ramakrishnan, and Shaidi (eds.), *Contributions to Automorphic Forms, Geometry and Arithmetic*, Johns Hopkins University Press.

Clozel, L. and Ullmo, E. (2005a) Equidistribution de mesures algébriques, *Compositio Math.* **141**, 1255–1309.

Clozel, L. and Ullmo, E. (2005b) Equidistribution de sous-variétés spéciales, *Ann. of Math. (2)* **161**, 1571–1588.

Cohen, P. (2005) Hyperbolic distribution problems on Siegel 3-folds and Hilbert modular varieties, *Duke Math. J.* **129**, 87–127.

Colmez, P. (1998) Sur la hauteur de Faltings des variétés abéliennesà multiplication complexe, *Compositio Math.* **111**, 359–368.

Dani, S. G. and Margulis, G. A. (1991) Asymptotic behaviour of trajectories of unipotent flows on homogeneous spaces, *Proc. Indian Acad. Sci. Math. Sci.* **101**, 1–17.

David, S. and Philippon, P. (2000) Sous-variétés de torsion des variétés semi-abéliennes, *C. R. Acad. Sci. Paris Sér. I Math.* **331**, 587–592.

Deligne, P. (1971) Travaux de Shimura, In *Séminaire Bourbaki, Exposé 389*, Vol. 244 of *Lecture Notes in Math.*, pp. 123–165, Berlin, Springer-Verlag.

Deligne, P. (1979) Variétés de Shimura: interprétation modulaire et techniques de construction de modèles canoniques, In A. Borel and W. Casselman (eds.), *Automorphic Forms, Representations, and L-functions. Part II*, Vol. 33 of *Proc. of Symp. in Pure Math.*, pp. 247–290, Amer. Math. Soc.

Duke, W. (1988) Hyperbolic distribution problems and half integral weights Maass-forms, *Invent. Math.* **92**, 73–90.

Edixhoven, B. (2005) Special points on products of modular curves, *Duke Math. J.* **126**, 325–348.

Edixhoven, B. and Yafaev, A. (2003) Subvarieties of Shimura varieties, *Ann. Math. (2)* **157**, 621–645.

Eskin, A. and Oh, H. (2006) Ergodic theoretic proof of equidistribution of Hecke points, *Ergodic Theory and Dynam. Systems*, to appear.

Faltings, G. (1984a) Arithmetic Varieties and Rigidity, In *Séminaire de Théorie des nombres de Paris*, Boston, MA, Birkhäuser.

Faltings, G. (1984b) Finiteness theorems for Abelian varieties over number fields, In *Arithmetic geometry*, Storrs, CO, 1984, pp. 9–27, New York, Springer.

Faltings, G. (1991) Diophantine approximation on Abelian varieties, *Ann. of Math. (2)* **133**, 549–576.

Hindry, M. (1988) Autour d'une conjecture de Serge Lang, *Invent. Math.* **94**, 575–603.

Hrushowski, E. (2001) The Manin–Mumford conjecture and the model theory of difference field, *Ann. Pure Appl. Logic* **112**, 43–115.

Iwaniec, H. (1995) Introduction to the spectral theory of automorphic forms, *Rev. Mat. Iberoamericana*.

Kim, H. and Sarnak, P. (2003) Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2. Appendix 2. Refined estimates towards the Ramanujan conjecture, *J. Amer. Math. Soc.* **16**, 139–183.

Milne, J. S. (2006) Introduction to Shimura varieties, www.jmilne.org/math/.

Milne, J. S. and Shih, K.-Y. (1982) Conjugates of Shimura varieties, In *Hodge cycles, motives, and Shimura varieties*, Vol. 900 of *Lecture Notes in Mathematics*, pp. 280–356, Berlin, Springer Verlag.

Moonen, B. (1998) Linearity properties of Shimura varieties. I, *J. Algebraic Geom.* **7**, 539–567.

Mozes, S. and Shah, N. (1995) On the space of ergodic invariant measures of unipotent flows, *Ergodic Theory Dynamic Systems* **15**, 149–159.

Pink, R. (2005) A combination of the conjectures of Mordell-Lang and André–Oort, In F. Bogomolov and Y. Tschinkel (eds.), *Geometric Methods in Algebra and Number Theory*, Vol. 253 of *Progress in Mathematics*, pp. 251–282, Birkhäuser.

Pink, R. and Roessler, D. (2004) On ψ-invariant subvarieties of semiAbelian varieties and the Manin–Mumford conjecture, *J. Algebraic Geom.* **13**, 771–798.

Platonov, V. and Rapinchuk, A. (1994) *Algebraic groups and number theory*, Vol. 139 of *Pure Appl. Math.*, Boston, MA, Academic Press Inc.

Ratner, M. (1991a) On Raghunathan's measure conjecture, *Ann. Math.* **134**, 545–607.

Ratner, M. (1991b) Raghunathan's topological conjecture and distribution of unipotent flows, *Duke Math. J.* **63**, 235–280.

Raynaud, M. (1988) Sous-variétés d'une variété abélienne et points de torsion, In *Arithmetic and Geometry. Vol. I*, Vol. 35 of *Progr. Math.*, pp. 327–352, Boston, MA, Birkhäuser.

Raynaud, R. (1983) Courbes sur une variété abélienne et points de torsion, *Invent. Math.* **71**, 207–223.

Szpiro, L., Ullmo, E., and Zhang, S. (1997) Equidistribution des petits points, *Invent. Math.* **127**, 337–347.

Ullmo, E. (1998) Positivité et discrètion des points algébriques des courbes, *Ann. of Math. (2)* **147**, 167–179.

Ullmo, E. (2002) Théorie ergodique et géométrie algébrique, In *Proceedings of the International Congress of Mathematicians*, Vol. 2, pp. 197–206, Higher Education Press.

Ullmo, E. (2006) Equidistribution de sous-variétés spéciales. II, *J. Reine Angew. Math.*, to appear.

Van der Geer, G. (1988) *Hilbert modular surfaces*, Vol. 16 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3)*, Berlin, Springer-Verlag.

Venkatesh, A. (2005) Sparse equidistribution problems, period bounds, and subconvexity, preprint.

Waldspurger, J.-L. (1985) Sur les valeurs de certaines fonctions L automorphes en leur centre de symétrie, *Compositio Math.* **54**, 173–242.

Wielonsky, F. (1985) Séries d'Eisenstein, intégrales toroïdales et une formule de Hecke, *Enseign. Math. (2)* **31**, 93–135.

Yafaev, A. (2006) A conjecture of Yves André, *Duke Math. J.*, to appear.

Zhang, S. (1998) Equidistribution of small points on Abelian varieties, *Ann. of Math. (2)* **147**, 159–165.

Zhang, S. (2005) Equidistribution of CM points on Quaternion Shimura Varieties, *Internat. Math. Res. Notices* **59**, 3657–3689.

ANALYTIC METHODS FOR THE DISTRIBUTION OF RATIONAL POINTS ON ALGEBRAIC VARIETIES

D. R. Heath-Brown
Oxford University

1. Introduction to the Hardy–Littlewood Circle Method

The most important analytic method for handling equidistribution questions about rational points on algebraic varieties is undoubtedly the Hardy–Littlewood circle method. There are a number of good texts available on the circle method, but the reader may particularly wish to study the books (Davenport, 2005) and (Vaughan, 1997). In this lecture we shall consider an irreducible form $F(X_1,\ldots,X_n) \in \mathbb{Z}[X_1,\ldots,X_n]$ of degree d which defines a hypersurface $F = 0$ in $\mathbb{P}^{n-1}$. The fundamental questions will be: are there any rational points on the hypersurface? If so, are they equidistributed in a suitable sense? We shall address these issues by looking at integer points on the affine cone. Thus to ask about equidistribution with respect to measures on $\mathbb{P}^{n-1}(\mathbb{R})$, for example, we may take a small box

$$\mathcal{R} := \prod_{i=1}^{n} (\kappa_i, \lambda_i] \subset \mathbb{R}^n,$$

in which

$$\kappa_i\lambda_i > 0, \quad (1 \le i \le n),$$

and examine the asymptotic behaviour of

$$N_{\mathcal{R}}(B) := \#\{\mathbf{x} \in \mathbb{Z}^n \cap B\mathcal{R} \colon F(\mathbf{x}) = 0\}$$

as $B \to \infty$. Here $B\mathcal{R} := \{\mathbf{x} \colon B^{-1}\mathbf{x} \in \mathcal{R}\}$. If we were to take a number of small boxes $\mathcal{R}_i$, of equal volumes, and we discovered that $N_{\mathcal{R}_i}(B)$ were asymptotically equal for each box, then we could say that the solutions of $F(\mathbf{x}) = 0$ were equidistributed amongst the boxes. Of course only boxes containing real solutions of $F(\mathbf{x}) = 0$ can feature, and indeed one finds in practice that points on the real locus $F = 0$ need to be appropriately weighted in order to

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 139–168.

achieve equidistribution. In other words what one attempts to establish is the equidistribution of integer solutions to $F(\mathbf{x}) = 0$, with respect to a suitable measure, on the real locus $F(\mathbf{x}) = 0$, and results on the behaviour of $N_{\mathcal{R}}(B)$ should be interpreted in this light.

Notice that the function $N_{\mathcal{R}}(B)$ may count a given projective point several times, with $\mathbf{x}$ being different scalar multiples of the same vector. However it is usually a trivial matter to adjust for this. In suitable circumstances one finds that

$$N_{\mathcal{R}}(B) \sim c(\mathcal{R})B^{n-d}$$

as $B \to \infty$, for a certain constant $c(\mathcal{R})$, and the equidistribution problem then becomes one of the behaviour of $c(\mathcal{R})$ as the box $\mathcal{R}$ varies.

The second aspect of equidistribution which we shall consider concerns the distribution of rational points within the p-adics, or more generally in the adèles. For this it is sufficient to select any vector $\mathbf{a} \in \mathbb{Z}^n$ and any modulus $m \in \mathbb{N}$ such that

$$(m, a_1, \ldots, a_n) = 1,$$

and investigate the modified counting function

$$N_{\mathcal{R}}(B; m, \mathbf{a}) := \#\{\mathbf{x} \in \mathbb{Z}^n \cap B\mathcal{R}\colon \mathbf{x} \equiv \mathbf{a} \pmod{m}, F(\mathbf{x}) = 0\}.$$

If m is composed of high powers of sufficiently many primes, then the behaviour of $N_{\mathcal{R}}(B; m, \mathbf{a})$ as both $\mathcal{R}$ and $\mathbf{a}$ vary is enough to describe completely the adèlic distribution of rational points on $F = 0$. Indeed the relative size of $N_{\mathcal{R}}(B; m, \mathbf{a})$, as $\mathbf{a}$ varies, describes the equidistribution of rational points, or lack of it, within the adèles.

Before proceeding further we should point out that although we have taken the groundfield as $\mathbb{Q}$ above, it is quite possible to consider problems over an arbitrary number field. The reader is referred to the book (Wang, 1991) for the circle method in this context.

Just what asymptotic behaviour should we expect for $N_{\mathcal{R}}(B)$? Since F is a form of degree d, we will have $F(\mathbf{x}) = O(B^d)$ for $\mathbf{x} \in B\mathcal{R}$. Hence the 'probability' that a particular integer value of $F(\mathbf{x})$ vanishes should be around B^{-d}. The number of integer vectors $\mathbf{x}$ in the region $B\mathcal{R}$ is of order B^n, so one might suppose that $N_{\mathcal{R}}(B)$ is likely to be of order precisely B^{n-d}. Of course this can only be meaningful when $n > d$. Moreover it can happen that the region $B\mathcal{R}$ does not even contain a real solution $\mathbf{x} \in \mathbb{R}^n$ of the equation $F(\mathbf{x}) = 0$. None the less we can state the following guiding principle.

HEURISTIC PRINCIPLE. *When $n > d$ we have*

$$B^{n-d} \ll N_{\mathcal{R}}(B) \ll B^{n-d}$$

unless there is a reason why not!

We should perhaps explain here how the Vinogradov $\ll$ notation is used. The statement $f \ll g$ means that there exists a positive "constant" c, such that $|f| \leq cg$ for all values under consideration. It follows that the statement $0 \ll g$ is somewhat uninformative, holding for any non-negative g! When we say that c is "constant" we mean that it does not depend on any variable, but may depend on parameters which are regarded as fixed. Thus, in our heuristic principle the implied constants are independent of B, but might depend on $F, \mathcal{R}, n$ and d.

The starting point for the Hardy–Littlewood circle method is the generating function

$$S(\alpha; B\mathcal{R}) = S(\alpha) := \sum_{\mathbf{x} \in \mathbb{Z}^n \cap B\mathcal{R}} \exp\{2\pi i \alpha F(\mathbf{x})\},$$

in which we take α to be a real variable, usually on $[0, 1]$. The function $\exp\{2\pi i\alpha\}$ occurs so frequently that we introduce the standard notation $e(\alpha) := \exp\{2\pi i\alpha\}$. Thus we may think of $e(.)$ as being a character on $\mathbb{R}^+/\mathbb{Z}^+$. We now have the key fact that if $n \in \mathbb{Z}$ then

$$\int_0^1 e(\alpha n)\, d\alpha = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases} \tag{1}$$

Applying this to the generating function $S(\alpha)$ produces the formula

$$N_{\mathcal{R}}(B) = \int_0^1 S(\alpha)\, d\alpha. \tag{2}$$

We may of course replace the range of integration by any interval of length 1.

What should we expect about the size of $S(\alpha)$? Let us suppose that the intervals $[\kappa_i, \lambda_i]$ defining $\mathcal{R}$ satisfy $-1 \leq \kappa_i < \lambda_i \leq 1$. This can always be achieved by re-scaling. Moreover let us set

$$V(\mathcal{R}) := \prod_{i=1}^{n} (\lambda_i - \kappa_i). \tag{3}$$

Then if we write $\mathcal{N}(B)$ for the number of integer vectors in $B\mathcal{R}$ we see that

$$\mathcal{N}(B) = \{V(\mathcal{R}) + o(1)\} B^n.$$

Now if $0 < \alpha < 1$ we might expect the numbers $e(\alpha F(\mathbf{x}))$ to be randomly scattered around the unit circle. Indeed if they are sufficiently random we might expect, from the central limit theorem, that $S(\alpha)$ should have a normal distribution in some sense. This leads us to expect that $S(\alpha)$ should usually be of order not much more than $B^{n/2}$. In other words, we expect that the terms in $S(\alpha)$ cancel to exactly the extent predicted by crude probabilistic arguments.

In contrast, when α is near to 0 (or to 1) we certainly can not expect such cancellation. To make this precise, let us write $||F||$ for the sum of the moduli of the coefficients of F, so that $|F(\mathbf{x})| \le ||F||B^d$ for $\mathbf{x} \in B\mathcal{R}$. Then if

$$\alpha \in \mathcal{J} := [-(8||F||B^d)^{-1}, (8||F||B^d)^{-1}],$$

say, we find that $|\alpha F(\mathbf{x})| \le \frac{1}{8}$ for $\mathbf{x} \in B\mathcal{R}$. Thus

$$\frac{1}{\sqrt{2}} = \cos\left(\frac{\pi}{4}\right) \le \Re\{e(\alpha F(\mathbf{x}))\} \le 1,$$

and hence

$$\frac{1}{\sqrt{2}}\mathcal{N}(B) \le \Re\{S(\alpha)\} \le \mathcal{N}(B).$$

Thus if we write

$$N_{\mathcal{R}}(B; \mathcal{J}) = \int_{\mathcal{J}} S(\alpha)\, d\alpha$$

for that part of the integral (2) in which α is suitably small, we find that

$$\{c_1 + o(1)\}B^{n-d} \le N_{\mathcal{R}}(B; \mathcal{J}) \le \{c_2 + o(1)\}B^{n-d}, \tag{4}$$

with $c_1 = V(\mathcal{R})/(4||F||\sqrt{2})$ and $c_2 = V(\mathcal{R})/(4||F||)$.

The precise values of c_1 and c_2 are unimportant, but there are several significant observations to be made. Firstly we see that the integration for α near 0 produces a term which is of exactly the order B^{n-d} we previously predicted heuristically. Conversely, if $N_{\mathcal{R}}(B)$ does not behave as the naive heuristics predict, so that it is not of exact order B^{n-d}, then the circle method is almost certainly doomed to fail. Thirdly, in an ideal world we may hope that the range in which α is not close to the endpoints of the interval $[0, 1]$ will only make a contribution of order about $B^{n/2}$. Of course, this will be smaller than the expected main term when $n > 2d$. Thus we might hope to prove that $N_{\mathcal{R}}(B)$ is of order B^{n-d}, provided that $n > 2d$. The final point is more technical, and is a converse to the third. In view of the central limit theorem we expect that $|S(\alpha)|$ is usually exactly of order $B^{n/2}$. Thus we cannot hope to prove results for the case $n \le 2d$ unless there is demonstrable cancellation in the values of $S(\alpha)$ as we average over α.

In any event, the basic message is that the method can only work, in its simplest form, if the number of variables exceeds $2d$. Moreover, in order to succeed we must prove that the sum $S(\alpha)$ cancels in the way that the central limit theorem predicts.

It is now time to admit that our initial distinction between values α near 0 or 1, and values in the remainder of the range was far too simple. It is quite possible that the coefficients of F are all even, for example, in which case we will have the same behaviour around $\alpha = \frac{1}{2}$ as we previously saw near $\alpha = 0$. More generally, it will often be the case that the values of $F(\mathbf{x})$ are not uniformly distributed to some particular modulus q, in which case $S(\alpha)$ may be large when α is close to a fraction of the form a/q. To make this more precise, we define

$$S_q(a) := \sum_{\mathbf{x} \,(\mathrm{mod}\, q)} e\left(\frac{a}{q}F(\mathbf{x})\right),$$

where the summation condition means that $\mathbf{x}$ runs over $\mathbb{Z}^n/q\mathbb{Z}^n$, or more concretely that $\mathbf{x}$ runs over integer vectors for which

$$0 \le x_i \le q-1, \quad (1 \le i \le n).$$

Clearly $e(aF(\mathbf{x})/q) = e(aF(\mathbf{y})/q)$ whenever $\mathbf{x} \equiv \mathbf{y} \,(\mathrm{mod}\, q)$, so that

$$S\left(\frac{a}{q}\right) = \sum_{\mathbf{x} \,(\mathrm{mod}\, q)} e\left(\frac{a}{q}F(\mathbf{x})\right) \#\{\mathbf{y} \in B\mathcal{R} \colon \mathbf{y} \equiv \mathbf{x} \,(\mathrm{mod}\, q)\}.$$

Moreover

$$\#\{y \in (Ba, Bb] \colon y \equiv x \,(\mathrm{mod}\, q)\} = (b-a)\frac{B}{q} + O(1),$$

whence

$$\#\{\mathbf{y} \in B\mathcal{R} \colon \mathbf{y} \equiv \mathbf{x} \,(\mathrm{mod}\, q)\} = V(\mathcal{R})\frac{B^n}{q^n} + O(B^{n-1}),$$

with $V(\mathcal{R})$ given by (3). It follows that

$$\begin{aligned} S\left(\frac{a}{q}\right) &= \sum_{\mathbf{x} \,(\mathrm{mod}\, q)} e\left(\frac{a}{q}F(\mathbf{x})\right) V(\mathcal{R})\frac{B^n}{q^n} + \sum_{\mathbf{x} \,(\mathrm{mod}\, q)} O(B^{n-1}) \\ &= V(\mathcal{R})\frac{B^n}{q^n}S_q(a) + O(q^n B^{n-1}). \end{aligned}$$

Thus, for any fixed fraction a/q, we see that if $S_q(a)$ is non-zero, then $S(a/q)$ will be of exact order B^n. Hence indeed $S(\alpha)$ will also be of exact order B^n when α is close to a/q.

It follows, via the analysis which led to (4), that values of α which are close to any fraction of small denominator may make contributions of the size predicted by our original heuristic argument. Thus it is appropriate to split the range $[0, 1]$ into two subsets, the "major arcs" in which α is close to a rational number with small denominator, in some precisely defined sense, and the remaining set, known as the "minor arcs." This terminology stems from the original formulation of the circle method, in which one integrated around a circle in the complex plane, instead of using the interval $[0, 1]$.

The reader will observe that everything we have described carries over immediately to $N_{\mathcal{R}}(B; m, \mathbf{a})$. All that is necessary is to work with the modified generating function

$$S(\alpha; B\mathcal{R}; m, \mathbf{a}) := \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap B\mathcal{R} \\ \mathbf{x}\equiv\mathbf{a}\ (\mathrm{mod}\, q)}} \exp\{2\pi i\alpha F(\mathbf{x})\}.$$

Having summarised the key features of the circle method we complete this lecture with a review of results. Statements in the literature are formulated in a variety of different ways. For example, some authors work only with $\mathcal{R} = [-1, 1]^n$; most do not consider the condition $\mathbf{x} \equiv \mathbf{a} \pmod{m}$; and sometimes one is content with establishing the existence of solutions rather than deriving an asymptotic formula for $N_{\mathcal{R}}(B)$. Thus the references we shall give will usually not contain the precise assertions we present below.

We begin with the case $d = 2$, so that F is a quadratic form. Here one has the following basic result, which we shall prove in the next two lectures.

THEOREM 1.1. *Let F be a diagonal quadratic form. If $n \geq 5$ there is a constant $c = c(\mathcal{R}; m, \mathbf{a})$ such that*

$$N_{\mathcal{R}}(B; m, \mathbf{a}) = cB^{n-2} + o(B^{n-2})$$

as $B \to \infty$.

We shall see a full interpretation of the constant c in the next lecture, but we should say at once that $c = 0$ if and only if the local conditions on $\mathbf{x}$ imply that $F(\mathbf{x})$ never vanishes.

Notice that the condition $n \geq 5$ corresponds to the inequality $n > 2d$ which we encountered earlier. In fact one can do a little better for quadratics, using a special variant of the circle method (see (Heath-Brown, 1996)), and show that

$$N_{\mathcal{R}}(B; m, \mathbf{a}) = cB^2 + o(B^2)$$

when $n = 4$ and the determinant of the quadratic form is not a square. However if the determinant is a square one gets a different behaviour

$$N_{\mathcal{R}}(B; m, \mathbf{a}) = cB^2 \log B + o(B^2 \log B).$$

One can even handle forms with $n = 3$, in which case one has

$$N_{\mathcal{R}}(B; m, \mathbf{a}) = cB \log B + o(B \log B).$$

We now turn to forms of degree 3. If $F(\mathbf{x})$ is an “additive” or “diagonal” form, of the shape

$$F(\mathbf{x}) = c_1x_1^3 + \cdots + c_nx_n^3,$$

the analysis can be pushed rather further than in the general case, and one has

$$N_{\mathcal{R}}(B; m, \mathbf{a}) = cB^{n-3} + o(B^{n-3})$$

as soon as $n \geq 8$. This follows from the methods of Vaughan (Vaughan, 1986). In this case we are not yet able to handle values of n as small as $n = 2d + 1$.

For non-diagonal cubic forms things are harder, but Hooley (Hooley, 1994) has shown essentially that

$$N_{\mathcal{R}}(B; m, \mathbf{a}) = cB^{n-3} + o(B^{n-3})$$

for $n \geq 9$, providing that F is non-singular. Finally, for arbitrary cubic forms our basic heuristic may fail completely, as the example

$$F(X_1, \ldots, X_n) = X_1^3 + X_2(X_3^2 + \cdots + X_n^2)$$

shows. Here one has $F(\mathbf{x}) = 0$ whenever $x_1 = x_2 = 0$. Thus one has $N_{\mathcal{R}}(B) \gg B^{n-2}$ for suitable regions $\mathcal{R}$.

For general degrees the diagonal case is again easier. Here one can show

$$N_{\mathcal{R}}(B; m, \mathbf{a}) = cB^{n-d} + o(B^{n-d})$$

for $n \geq n_1(d)$ with a value of $n_1(d)$ such that

$$n_1(d) \sim d^2 \log d$$

by the method of (Ford, 1995). For general non-singular forms one can establish the same asymptotic relation for $N_{\mathcal{R}}(B; m, \mathbf{a})$ by the method of (Birch, 1962), under the condition $n \geq n_2(d)$, where $n_2(d) = 1 + (d-1)2^d$.

The overall picture is thus that the circle method only works if the number of variables is sufficiently large in terms of the degree, where “sufficiently large” refers to a function which tends to infinity with d. None the less, where the method does work it provides a complete answer to the equidistribution question. For singular varieties the heuristic formula may fail altogether.

We end this lecture by remarking that the techniques described here can be extended to varieties defined by the vanishing of a system of forms $F_1(\mathbf{x}) = \cdots = F_k(\mathbf{x}) = 0$. All that is necessary is to define a multi-variable generating function

$$S(\alpha_1, \ldots, \alpha_k) := \sum_{\mathbf{x} \in \mathbb{Z}^n \cap B\mathcal{R}} \exp\{2\pi i(\alpha_1 F_1(\mathbf{x}) + \cdots + \alpha_k F_k(\mathbf{x}))\},$$

and integrate over $[0,1]^k$. Examples of results which can be proved in this way may be found in (Davenport and Lewis, 1969), and in (Birch, 1962) for general forms.

2. Major Arcs and Local Factors in the Hardy–Littlewood Circle Method

In this lecture we shall work with the integral

$$N_{\mathcal{R}}(B; m, \mathbf{a}) = \int_0^1 S(\alpha; B\mathcal{R}; m, \mathbf{a})\, d\alpha \tag{5}$$

analogous to (2), and consider the contribution coming from α in the "major arcs." For the time being we shall merely assume that $F(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ is a form of degree d in n variables. Later we shall have to impose further restrictions on F and d.

Clearly it is time to be more specific about what we mean by the "major arcs." We shall use a small positive parameter $\delta < \frac{1}{3}$, independent of B, which we shall specify later, see (11). It will be convenient to write $Q := B^\delta$. For each positive integer $q \leq Q$, and each non-negative integer $a < q$ such that $(a, q) = 1$, we proceed to define the interval

$$I(a, q) := \left[\frac{a}{q} - \frac{Q}{B^d}, \frac{a}{q} + \frac{Q}{B^d}\right].$$

We now observe that these various intervals are disjoint providing that B is large enough, as we henceforth assume. For if $a/q \neq a'/q'$ then

$$\left|\frac{a}{q} - \frac{a'}{q'}\right| \geq \frac{1}{qq'} \geq \frac{1}{Q^2} > 2\frac{Q}{B^d}.$$

With this in mind we define the set of "major arcs" as

$$\mathfrak{M} = \bigcup_{q \leq Q} \bigcup_{\substack{0 \leq a < q \\ (a,q)=1}} I(a, q).$$

Strictly speaking we do not have $\mathfrak{M} \subseteq [0, 1]$ since the left endpoint of $I(0, 1)$ is negative. Thus we will replace the range of integration in (5) by the interval $[-Q/B^d, 1 - Q/B^d]$ to avoid this problem. We then define the "minor arcs" to be the set

$$\mathfrak{m} = \left[-\frac{Q}{B^d}, 1 - \frac{Q}{B^d}\right] \setminus \mathfrak{M}.$$

It should be pointed out at this stage that there is considerable flexibility in choosing the major arc set, and for other applications a rather different choice may be appropriate. Indeed in some cases one chooses the major arcs so as to comprise the entire interval $[0, 1]$. The key points to note about our choice (indeed any choice) for the set $\mathfrak{M}$, are firstly that the intervals $I(a, q)$ are longer than B^{-d} by a factor which tends to infinity; and secondly that one uses all denominators up to a bound Q which also tends to infinity. The reader will note that the description given in the first lecture fits neither of these conditions!

We now face the crucial task of approximating $S(\alpha) = S(\alpha; B\mathcal{R}; m, \mathbf{a})$ on the various intervals $I(a, q)$. We will certainly need more precise estimates than we obtained in the first lecture. It will be convenient to put $\theta = \alpha - a/q$. Then if $\mathbf{y} \equiv \mathbf{x} \pmod{q}$ we have

$$e(\alpha F(\mathbf{y})) = e\left(\frac{a}{q}F(\mathbf{y})\right)e(\theta F(\mathbf{y})) = e\left(\frac{a}{q}F(\mathbf{x})\right)e(\theta F(\mathbf{y})).$$

In fact we also have the congruence condition $\mathbf{x} \equiv \mathbf{a} \pmod{m}$ to contend with. The reader may prefer to take $m = 1$ in what follows, but we shall give the details for the general case. We have

$$\begin{aligned} S(\alpha; B\mathcal{R}; m, \mathbf{a}) &= \sum_{\substack{\mathbf{x} \pmod{mq} \\ \mathbf{x}\equiv\mathbf{a} \pmod{m}}} \sum_{\substack{\mathbf{y}\in B\mathcal{R} \\ \mathbf{y}\equiv\mathbf{x} \pmod{mq}}} e(\alpha F(\mathbf{y})) \\ &= \sum_{\substack{\mathbf{x} \pmod{mq} \\ \mathbf{x}\equiv\mathbf{a} \pmod{m}}} e\left(\frac{a}{q}F(\mathbf{x})\right) \sum_{\substack{\mathbf{y}\in B\mathcal{R} \\ \mathbf{y}\equiv\mathbf{x} \pmod{mq}}} e(\theta F(\mathbf{y})). \qquad (6) \end{aligned}$$

(Here, and elsewhere in these notes, we shall follow the convention that summations are over integers, and integer vectors, without explicitly writing $\mathbf{y} \in \mathbb{Z}^n \cap B\mathcal{R}$, for example.) The inner sum above contains an exponential in which $\theta F(\mathbf{y})$ is relatively small, of order $Q = B^{\delta}$, and smoothly varying. We can therefore hope to approximate the sum fairly well by the corresponding integral.

In fact we have

$$\sum_{\substack{\mathbf{y}\in B\mathcal{R} \\ \mathbf{y}\equiv\mathbf{x} \pmod{mq}}} e(\theta F(\mathbf{y})) = (mq)^{-n} I(B; \theta) + O(B^{n-1+2\delta}), \qquad (7)$$

where

$$I(B; \theta) := \int_{B\mathcal{R}} e(\theta F(\mathbf{z}))\, dz_1 \cdots dz_n. \qquad (8)$$

We should stress at this point that, in writing the above error term, we have taken the form F and the box $\mathcal{R}$, along with $m, \mathbf{a}$ and the parameter δ, to be fixed. Thus the error term is of the shape $\rho B^{n-1+2\delta}$, where ρ is bounded in terms of B, but not necessarily as a function of $F, \mathcal{R}, m$, $\mathbf{a}$ and δ. This remark will apply to all the error terms we shall meet henceforth.

The proof of (7) is somewhat technical, and might reasonably be skipped on a first reading. All that is important is the form of the main term, and the fact that the error term contains an exponent which is strictly less than n. However we present below the necessary argument.

For each relevant integer point $\mathbf{y}$ we define

$$C(\mathbf{y}) = \{\mathbf{z} \in \mathbb{R}^n : y_i < z_i \leq y_i + mq, \text{ for } 1 \leq i \leq n\}.$$

Then if $\mathbf{z} \in C(\mathbf{y})$ we have $F(\mathbf{z}) - F(\mathbf{y}) = O(qB^{d-1})$, since $|z_i - y_i| \leq mq \leq mQ \leq mB$ for each index i, and m is regarded as fixed. It follows that $\theta(F(\mathbf{z}) - F(\mathbf{y})) = O(QB^{-d}.qB^{d-1}) = O(Q^2B^{-1})$, whence $e(\theta F(\mathbf{z}) - \theta F(\mathbf{y})) = 1 + O(Q^2B^{-1})$ and thus

$$e(\theta F(\mathbf{z})) = e(\theta F(\mathbf{y})) + O(B^{-1+2\delta}).$$

If we integrate over $C(\mathbf{y})$ we may then deduce that

$$\int_{C(\mathbf{y})} e(\theta F(\mathbf{z}))\, dz_1 \cdots dz_n = (mq)^n e(\theta F(\mathbf{y})) + O((mq)^n B^{-1+2\delta}),$$

so that

$$e(\theta F(\mathbf{y})) = (mq)^{-n} \int_{C(\mathbf{y})} e(\theta F(\mathbf{z}))\, dz_1 \cdots dz_n + O(B^{-1+2\delta}).$$

We have constructed the cubes $C(\mathbf{y})$ so as to be disjoint. Moreover the box $B\mathcal{R}$ contains $O(B^n)$ integer points $\mathbf{y}$. Hence, when we sum over the various values of $\mathbf{y}$ we find that

$$\sum_{\substack{\mathbf{y} \in B\mathcal{R} \\ \mathbf{y} \equiv \mathbf{x} \pmod{mq}}} e(\theta F(\mathbf{y})) = (mq)^{-n} \int_C e(\theta F(\mathbf{z}))\, dz_1 \cdots dz_n + O(B^{n-1+2\delta}),$$

where C is the union of the cubes $C(\mathbf{y})$. If we have a point $\mathbf{z}$ which belongs to one of $B\mathcal{R}$ or C but not to the other, then $\mathbf{z}$ must lie at a distance at most $O(mq) = O(q)$ from the boundary of $B\mathcal{R}$. Since $q \leq B$ it follows that

$$\int_C e(\theta F(\mathbf{z}))\, dz_1 \cdots dz_n = \int_{B\mathcal{R}} e(\theta F(\mathbf{z}))\, dz_1 \cdots dz_n + O(qB^{n-1}).$$

On comparing our various estimates we therefore derive the required approximation (7).

We now return to the main thread of the argument by inserting (7) into (6) to deduce that

$$S(\alpha; B\mathcal{R}; m, \mathbf{a}) = \frac{S_q(a; m, \mathbf{a})}{(mq)^n} I(B; \theta) + O(q^n B^{n-1+2\delta}),$$

where we have defined

$$S_q(a; m, \mathbf{a}) := \sum_{\substack{\mathbf{x} \ \mathrm{mod}(mq) \\ \mathbf{x} \equiv \mathbf{a} \ (\mathrm{mod}\, m)}} e\left(\frac{a}{q} F(\mathbf{x})\right).$$

We proceed to integrate over $I(a, q)$, producing

$$\int_{I(a,q)} S(\alpha; B\mathcal{R}; m, \mathbf{a})\, d\alpha = \frac{S_q(a; m, \mathbf{a})}{(mq)^n} J(B) + O(Q^{n+1} B^{n-d-1+2\delta})$$

with

$$J(B) := \int_{-QB^{-d}}^{QB^{-d}} I(B, \theta)\, d\theta. \tag{9}$$

We can then sum for a coprime to q and for $q \le Q$ to yield

$$\begin{aligned} \int_{\mathfrak{M}} S(\alpha; B\mathcal{R}; m, \mathbf{a})\, d\alpha &= S(Q) J(B) + O(Q^{n+3} B^{n-d-1+2\delta}) \\ &= S(Q) J(B) + O(B^{n-d-1+(n+5)\delta}), \qquad (10) \end{aligned}$$

where

$$S(Q) := \sum_{q \le Q} \sum_{\substack{0 \le a < q \\ (a,q)=1}} \frac{S_q(a; m, \mathbf{a})}{(mq)^n}.$$

The error term here is $o(B^{n-d})$, which is satisfactory for Theorem 1.1, providing that we choose

$$\delta = \frac{1}{n+6}, \tag{11}$$

as we shall indeed do.

Thus to complete the treatment of the major arcs we need to say more about the sum $S(Q)$ and the integral $J(B)$. We begin with the latter. By substituting $\mathbf{z} = B\mathbf{w}$ in (8) and $\theta = B^{-d}\phi$ in (9) we find that

$$J(B) = B^{n-d} \int_{-Q}^{Q} \left\{ \int_{\mathcal{R}} e(\phi F(\mathbf{w}))\, dw_1 \cdots dw_n \right\} d\phi.$$

We now impose an additional technical requirement on the form F, namely that

$$F_1(w_1, \ldots, w_n) > 0 \ \forall \mathbf{w} \in \mathcal{R}, \tag{12}$$

where

$$F_1(w_1,\ldots,w_n) := \frac{\partial F(w_1,\ldots,w_n)}{\partial w_1}.$$

(With a little more effort one can analyse $J(B)$ satisfactorily under the more natural assumption that F is non-singular in $\mathcal{R}$.) We can then solve the equation $u = F(w_1,\ldots,w_n)$ for w_1, using the Implicit Function Theorem. This produces a function $w_1 = f(u,w_2,\ldots,w_n)$, say. We proceed to define

$$g(u) := \int \frac{dw_2\cdots dw_n}{F_1(f(u,w_2,\ldots,w_n),w_2,\ldots,w_n)},$$

where the integration is subject to $(w_1,\ldots,w_n) \in \mathcal{R}$. The function g is continuously differentiable, and has compact support. (The first of these two assertions requires more than a moment's thought.) With these definitions we have

$$\int_{\mathcal{R}} e(\phi F(\mathbf{w}))\,dw_1\cdots dw_n = \int_{-\infty}^{\infty} e(\phi u)g(u)\,du = \hat{g}(-\phi),$$

where $\hat{g}$ denotes the Fourier transform. It then follows that

$$\lim_{B\to\infty}\int_{-B^\delta}^{B^\delta}\left\{\int_{\mathcal{R}} e(\phi F(\mathbf{w}))\,dw_1\cdots dw_n\right\}d\phi = \lim_{T\to\infty}\int_{-T}^{T}\hat{g}(-\phi)\,d\phi = g(0),$$

whence

$$J(B) \sim B^{n-d}c_\infty \quad (B\to\infty) \tag{13}$$

where c_∞ is the so called "singular integral," given by

$$c_\infty := g(0) = \int_{\mathbf{w}\in\mathcal{R},F(\mathbf{w})=0}\frac{dw_2\cdots dw_n}{F_1(w_1,\ldots,w_n)},$$

with a certain amount of abuse of notation. The singular integral is to be regarded as the density of real points of $F = 0$ in the box $\mathcal{R}$. In fact one can show, subject only to the condition that F is non-singular in $\mathcal{R}$, that

$$c_\infty = \lim_{\varepsilon\downarrow 0}\frac{1}{2\varepsilon}\mathrm{Meas}\{\mathbf{w}\in\mathcal{R}\colon |F(\mathbf{w})| < \varepsilon\}.$$

We end our discussion of $J(B)$ by remarking that (13) shows in particular that $J(B) = O(B^{n-d})$.

We turn now to the sum $S(Q)$. Just as $J(B)$ reflects the behaviour of F with respect to the real valuation $|.|_\infty$, we shall see that $S(Q)$ contains the corresponding p-adic information. Our treatment will require the following lemma, which we will use to separate out the factors corresponding to individual primes.

LEMMA 2.1. *Let $q = rs$ and $m = jk$, with rj coprime to sk. Choose integers $\bar{r}$ and $\bar{s}$ such that $r\bar{r} \equiv 1 \pmod{sk}$ and $s\bar{s} \equiv 1 \pmod{rj}$. Then*

$$S_q(a; m, \mathbf{a}) = S_r(a\bar{s}; j, \mathbf{a})S_s(a\bar{r}; k, \mathbf{a}).$$

For the purposes of the proof we choose integers $\bar{j}$ and $\bar{k}$ with $j\bar{j} \equiv 1 \pmod{sk}$ and $k\bar{k} \equiv 1 \pmod{rj}$. One can then verify that if $\mathbf{u}$ runs over vectors modulo rj with $\mathbf{u} \equiv \mathbf{a} \pmod{j}$, and $\mathbf{v}$ runs over vectors modulo sk with $\mathbf{v} \equiv \mathbf{a} \pmod{k}$, then

$$\mathbf{w} := sk\bar{s}\bar{k}\mathbf{u} + rj\bar{r}\bar{j}\mathbf{v}$$

runs over vectors modulo qm with $\mathbf{w} \equiv \mathbf{a} \pmod{m}$. Moreover we have

$$F(\mathbf{w}) \equiv (sk\bar{s}\bar{k})^d F(\mathbf{u}) + (rj\bar{r}\bar{j})^d F(\mathbf{v}) \equiv s\bar{s}F(\mathbf{u}) + r\bar{r}F(\mathbf{v}) \pmod{q},$$

whence

$$e\left(\frac{a}{q}F(\mathbf{w})\right) = e\left(\frac{a\bar{s}}{r}F(\mathbf{u})\right)e\left(\frac{a\bar{r}}{s}F(\mathbf{v})\right).$$

The lemma then follows on summing over $\mathbf{u}$ and $\mathbf{v}$.

We now examine the sum

$$S(q, m, \mathbf{a}) := \sum_{\substack{0 \le a < q \\ (a,q)=1}} S_q(a; m, \mathbf{a})$$

occurring in the definition of $S(Q)$. With the above notation we note that if b and c run over distinct residue classes coprime to r and s respectively, then $a = sb + rc$ will run over distinct residue classes coprime to rs. Then Lemma 2.1 implies that

$$\begin{aligned} S(q, m, \mathbf{a}) &= \sum_{\substack{0 \le b < r \\ (b,r)=1}} \sum_{\substack{s0 \le c < s \\ (c,s)=1}} S_q(sb + rc; m, \mathbf{a}) \\ &= \sum_{\substack{0 \le b < r \\ (b,r)=1}} \sum_{\substack{0 \le c < s \\ (c,s)=1}} S_r((sb + rc)\bar{s}; j, \mathbf{a})S_s((sb + rc)\bar{r}; k, \mathbf{a}) \\ &= \sum_{\substack{0 \le b < r \\ (b,r)=1}} \sum_{\substack{0 \le c < s \\ (c,s)=1}} S_r(b; j, \mathbf{a})S_s(c; k, \mathbf{a}) \\ &= S(r, j, \mathbf{a})S(s, k, \mathbf{a}). \end{aligned}$$

It follows that if $q = \prod p^e$ and $m = \prod p^f$ then we have

$$S(q, m, \mathbf{a}) = \prod_p S(p^e, p^f, \mathbf{a}). \tag{14}$$

Thus we can think of $S(Q)$ as being a partial sum corresponding to an Euler product. In what follows we shall follow the above notation, writing $f = \nu_p(m)$ for each prime p, for example.

Our next goal will be to replace $S(Q)$ by the corresponding infinite sum. This will require estimates for the size of $S(q, m, \mathbf{a})$ which may be derived by the methods in the next lecture. Thus, for the time being we content ourselves with working on the following hypothesis.

HYPOTHESIS. *There is a real number $\eta > 0$, depending only on n and d, such that*

$$S_q(a; m, \mathbf{a}) \ll q^{n-2-\eta}$$

whenever a and q are coprime.

We should remark that if q is prime, and if F is non-singular modulo q, then one has

$$S_q(a; m, \mathbf{a}) \ll q^{n/2},$$

by Deligne's Riemann Hypothesis for varieties over finite fields (Deligne, 1980). When $n \geq 5$ this result yields good information for composite moduli q which have no square factor, by Lemma 2.1. However there are serious problems in extending such estimates to prime powers.

Under the hypothesis above we see that $S(q, m, \mathbf{a}) \ll q^{n-1-\eta}$, so that

$$\sum_{q=1}^{\infty} q^{-n} S(q, m, \mathbf{a}) \tag{15}$$

is absolutely convergent. It then follows from (14) that

$$\sum_{q=1}^{\infty} q^{-n} S(q, m, \mathbf{a}) = \prod_p c_p,$$

where

$$c_p := \sum_{e=0}^{\infty} p^{-en} S(p^e, p^f, \mathbf{a})$$

with $f = \nu_p(m)$. Indeed we will have

$$\sum_{q>Q} q^{-n} |S(q, m, \mathbf{a})| \ll Q^{-\eta},$$

whence

$$S(Q) = m^{-n} \prod_p c_p + O(Q^{-\eta}).$$

We can now combine this with (10) and (13) to conclude that

$$\int_{\mathfrak{M}} S(\alpha; B\mathcal{R}; m, \mathbf{a})\, d\alpha \sim B^{n-d} m^{-n} c_\infty \prod_p c_p \quad (B \to \infty), \tag{16}$$

subject to the hypothesis above on the size of $S_q(a; m, \mathbf{a})$, and on our previous assumption (12).

The sum (15) is known as the "singular series." The original formulation of the circle method used a function of a complex variable integrated around a circle, and there was a singularity at each point $e(a/q)$ with $q \le Q$. Thus the series arose from the contributions of these various singularities.

We close this lecture by giving an important interpretation of the factors c_p, which depends on the following lemma.

LEMMA 2.2. *Define*

$$N(p^e, p^f, \mathbf{a}) := \#\{\mathbf{x} \pmod{p^{e+f}} : \mathbf{x} \equiv \mathbf{a} \pmod{p^f},\, p^e | F(\mathbf{x})\}.$$

Then

$$S(p^e, p^f, \mathbf{a}) = p^e N(p^e, p^f, \mathbf{a}) - p^{n+e-1} N(p^{e-1}, p^f, \mathbf{a})$$

for $e \ge 1$.

By definition we have

$$S(p^e, p^f, \mathbf{a}) = \sum_{a=0}^{p^e-1} S_{p^e}(a; p^f, \mathbf{a}) - \sum_{a=0}^{p^{e-1}-1} S_{p^e}(ap; p^f, \mathbf{a}).$$

Moreover we have

$$S_{p^e}(ap; p^f, \mathbf{a}) = p^n S_{p^{e-1}}(a; p^f, \mathbf{a}),$$

since

$$e\left(\frac{ap}{p^e} F(\mathbf{x})\right) = e\left(\frac{a}{p^{e-1}} F(\mathbf{y})\right)$$

for every $\mathbf{x}$ such that $\mathbf{x} \equiv \mathbf{y} \pmod{p^{e+f-1}}$. For the proof of the lemma it now suffices to observe that

$$N(p^e, p^f, \mathbf{a}) = \sum_{a=0}^{p^e-1} S_{p^e}(a, p^f, \mathbf{a}),$$

since

$$\sum_{a=0}^{q} e\left(\frac{a}{q} w\right) = \begin{cases} q, & q \mid w, \\ 0, & q \nmid w. \end{cases}$$

We proceed to consider the behaviour of $N(p^e, p^f, \mathbf{a})$ as e tends to infinity. An induction argument based on Lemma 2.2 shows that

$$\sum_{e=0}^{E} p^{-en} S(p^e, p^f, \mathbf{a}) = p^{-E(n-1)} N(p^E, p^f, \mathbf{a})$$

for every $E \geq 0$, whence

$$c_p = \lim_{e \to \infty} p^{-e(n-1)} N(p^e, p^f, \mathbf{a}).$$

We may therefore think of c_p as being the density of points on $F = 0$ in $\mathbb{A}^n(\mathbb{Z}_p)$, subject to the condition $\mathbf{x} \equiv \mathbf{a} \pmod{m}$. Thus the product $c_\infty \prod_p c_p$ in (16) can be interpreted as a product of local densities, one for each place of $\mathbb{Q}$. It can be shown that $c_p > 0$ if and only if the variety $F = 0$ has a non-singular p-adic point $\mathbf{x}$ with $\mathbf{x} \equiv \mathbf{a} \pmod{m}$.

Now, suppose that the circle method succeeds in showing that

$$N_{\mathcal{R}}(B; m, \mathbf{a}) \sim B^{n-d} m^{-n} c_\infty \prod_p c_p \quad (B \to \infty)$$

for every $\mathbf{a}$ and m. We can then show that the rational points on $F = 0$ are dense in the adèlic points, providing that the non-singular adèlic points are dense in the set of all adèlic solutions of $F = 0$. To see this, let $\mathcal{N}$ be an adèlic neighbourhood containing a point on $F = 0$. According to our assumption, $\mathcal{N}$ will contain a non-singular adèlic point $a = (\mathbf{a}_\infty, \mathbf{a}_2, \mathbf{a}_3, \ldots)$ on $F = 0$. It follows that there exists $\varepsilon > 0$ and p_0 such that $\mathcal{N}$ contains the set of all adèles $(\mathbf{b}_\infty, \mathbf{b}_2, \mathbf{b}_3, \ldots)$ satisfying $|\mathbf{b}_\infty - \mathbf{a}_\infty| < \varepsilon$ and $|\mathbf{b}_p - \mathbf{a}_p| < \varepsilon$ for $p \leq p_0$. Thus if $\mathbf{x} \in \mathbb{Z}^n$ we will have $B^{-1}\mathbf{x} \in \mathcal{N}$ providing that $\mathbf{x} \in B\mathcal{R}$ and $\mathbf{x} \equiv B\mathbf{a} \pmod{m}$, for a suitable box $\mathcal{R}$ and modulus m. Since $\mathbf{a}$ is non-singular, the circle method provides an asymptotic formula for the number of relevant rational points of the form $B^{-1}\mathbf{x}$, in which the constant $c_\infty \prod_p c_p$ is strictly positive. The claim then follows.

3. The Minor Arcs in the Hardy–Littlewood Circle Method

In this third lecture we shall consider the minor arc integral

$$\int_{\mathfrak{m}} S(\alpha; B\mathcal{R}; m, \mathbf{a})\, d\alpha.$$

Our primary goal is to show that this is $o(B^{n-d})$. However we also have a secondary goal, which is to establish the hypothesis used in the previous lecture to establish the convergence of the singular series. Up to now we have needed only mild conditions on the form F and its degree d, but unfortunately we shall now impose rather severe restrictions.

We shall begin by assuming that F is a diagonal form, so that it is of the shape

$$F(\mathbf{x}) = c_1x_1^d + \cdots + c_nx_n^d$$

for certain non-zero integer coefficients c_j. Under this assumption the generating function factorizes as

$$S(\alpha; B\mathcal{R}; m, \mathbf{a}) = \prod_{j=1}^{n} S_j(\alpha; B; m, a_j), \tag{17}$$

where

$$S_j(\alpha; B; m, a_j) := \sum_{\substack{x\in\mathbb{Z}\cap[B\kappa_j, B\lambda_j]\\ x\equiv a_j \pmod{m}}} e(\alpha c_j x^d).$$

For this factorisation to take place it is important that $\mathcal{R}$ is a box, as well as having F diagonal.

We are now able to consider a one-variable exponential sum, rather than an n-variable sum, and this produces a considerable simplification. We begin by using Hölder's inequality to show that

$$\begin{aligned}
&\left|\int_{\mathfrak{m}} S(\alpha; B\mathcal{R}; m, \mathbf{a})\, d\alpha\right| \\
&\qquad \le \int_{\mathfrak{m}} \prod_{j=1}^{n} |S_j(\alpha; B; m, a_j)|\, d\alpha \\
&\qquad \le \sup_{\alpha\in\mathfrak{m}} \prod_{j\le n-4} |S_j(\alpha; B; m, a_j)| \int_{\mathfrak{m}} \prod_{j=n-3}^{n} |S_j(\alpha; B; m, a_j)|\, d\alpha \\
&\qquad \le \sup_{\alpha\in\mathfrak{m}} \max_{j\le n-4} |S_j(\alpha; B; m, a_j)|^{n-4} \left\{\prod_{j=n-3}^{n} \int_{\mathfrak{m}} |S_j(\alpha; B; m, a_j)|^4\, d\alpha\right\}^{1/4} \\
&\qquad \le \sup_{\alpha\in\mathfrak{m}} \max_{j\le n-4} |S_j(\alpha; B; m, a_j)|^{n-4} \max_{n-3\le j\le n} \int_{\mathfrak{m}} |S_j(\alpha; B; m, a_j)|^4\, d\alpha \\
&\qquad \le \sup_{\alpha\in\mathfrak{m}} \max_{j\le n-4} |S_j(\alpha; B; m, a_j)|^{n-4} \max_{n-3\le j\le n} \int_0^1 |S_j(\alpha; B; m, a_j)|^4\, d\alpha.
\end{aligned}$$

It therefore follows that there are indices j and k for which

$$\left|\int_{\mathfrak{m}} S(\alpha; B\mathcal{R}; m, \mathbf{a})\, d\alpha\right| \le \sup_{\alpha\in\mathfrak{m}} |S_j(\alpha; B; m, a_j)|^{n-4} \int_0^1 |S_k(\alpha; B; m, a_k)|^4\, d\alpha. \tag{18}$$

We first examine the integral. We have

$$|S_k(\alpha; B; m, a_k)|^4 = \sum_{x_1,x_2,x_3,x_4} e(\alpha c_j(x_1^d + x_2^d - x_3^d + x_4^d)),$$

where the variables are subject to $x_i \in [B\kappa_k, B\lambda_k]$ and $x_i \equiv a_k \pmod m$. On applying (1) we therefore find that

$$\int_0^1 |S_k(\alpha; B; m, a_k)|^4 \, d\alpha = \#\{(x_1, x_2, x_3, x_4)\colon x_1^d + x_2^d = x_3^d + x_4^d\}, \qquad (19)$$

with the same constraints on the variables x_i as before. Since $-1 \le \kappa_k < \lambda_k \le 1$ we deduce that

$$\int_0^1 |S_k(\alpha; B; m, a_k)|^4 d\alpha \le \#\{(x_1, x_2, x_3, x_4) \in \mathbb{Z}^4 \cap [-B, B]^4\colon x_1^d + x_2^d = x_3^d + x_4^d\}.$$

We can estimate this quantity relatively easily, using the following well-known bound.

LEMMA 3.1. *For any $\varepsilon > 0$ there is a constant $c(\varepsilon)$ such that*

$$\#\{m \in \mathbb{N}\colon m \mid N\} \le c(\varepsilon) N^{\varepsilon}$$

for every positive integer N.

When $x_1^d > x_3^d$ we apply this to $N = x_1^d - x_3^d$, observing that $m = |x_2 - x_4|$ will be a positive integer divisor of N. Moreover we may note that the values of $N = x_4^d - x_2^d$ and $x_4 - x_2$ determine at most $d - 1$ possible values of x_2 and x_4. It follows that each pair x_1, x_3 with $x_1^d > x_3^d$ leads to $O(B^{d\varepsilon})$ quadruples (x_1, x_2, x_3, x_4). The case $x_1^d < x_3^d$ is of course similar, so that there are a total of $O(B^{2+d\varepsilon})$ solutions with $x_1^d \ne x_3^d$. Finally we observe that if $x_1^d = x_3^d$ then $x_2^d = x_4^d$. Thus we get $O(B^2)$ solutions in total this way. It therefore follows that

$$\int_0^1 |S_k(\alpha; B; m, a_k)|^4 \, d\alpha \ll B^{2+d\varepsilon} \qquad (20)$$

for any fixed $\varepsilon > 0$. One of the remarkable aspects of the circle method is the way in which the proof of the asymptotic behaviour of $N_{\mathcal{R}}(B)$ is made to depend on the behaviour of the counting function for a completely different variety, namely $x_1^d + x_2^d = x_3^d + x_4^d$. Indeed for this latter variety a crude upper bound is sufficient.

It remains to bound an individual value of $S_j(\alpha; B; m, a_j)$, so as to deal with the supremum on the right hand-side of (18). Here we shall handle only the case $d = 2$. Thus for the rest of this lecture we shall assume that $d = 2$, with the specific goal of proving Theorem 1.1. The procedure we use dates back to the work of Weyl (Weyl, 1916). We shall give an exposition here that is specific to our situation. For convenience of notation we put

$$\beta := \alpha c_j, \quad \kappa := \kappa_j, \quad \lambda := \lambda_j, \quad \text{and} \quad a := a_j,$$

so that

$$S_j(\alpha; B; m, a_j) = \sum_{\substack{\kappa B < x \le \lambda B \\ x \equiv a \ (\mathrm{mod}\, m)}} e(\beta x^2).$$

The argument begins with the expansion

$$|S_j(\alpha; B; m, a_j)|^2 = \sum_{x,y} e(\beta(x^2 - y^2)).$$

We then set $x = y + mz$, so that z runs over integer values in the range $|z| \le (\lambda - \kappa)B/m$. Thus in particular we have $|z| \le 2B$. For each value of z the conditions on y become

$$y + mz \in (\kappa B, \lambda B], \quad y \in (\kappa B, \lambda B], \quad \text{and} \quad y \equiv a \ (\mathrm{mod}\, m). \tag{21}$$

We now substitute $y = a + mw$ and set

$$\mu := \{\max(\kappa B - mz, \kappa B) - a\}/m, \quad \nu := \{\min(\lambda B - mz, \lambda B) - a\}/m,$$

so that the range (21) for y corresponds to the range $\mu < w \le \nu$ for w. The substitutions we have made yield

$$x^2 - y^2 = m^2 z^2 + 2amz + 2m^2 zw,$$

whence

$$\begin{aligned} |S_j(\alpha; B; m, a_j)|^2 &= \sum_{|z| \le 2B} e(\beta(m^2 z^2 + 2amz)) \sum_{\mu < w \le \nu} e(2\beta m^2 zw) \\ &\le \sum_{|z| \le 2B} \left| \sum_{\mu < w \le \nu} e(2\beta m^2 zw) \right|. \end{aligned}$$

Now in general, for any integers $W_1 < W_2$, and for any $\gamma \in \mathbb{R} \setminus \mathbb{Z}$, we have

$$\begin{aligned} \sum_{W_1 \le w \le W_2} e(\gamma w) &= \frac{e(\gamma(W_2 + 1)) - e(\gamma W_1)}{e(\gamma) - 1} \\ &= \frac{e(\gamma(W_2 + 1/2)) - e(\gamma(W_1 - 1/2))}{2i \sin(\pi\gamma)}, \end{aligned}$$

whence

$$\left|\sum_{W_1\le w\le W_2} e(\gamma w)\right| \le \frac{1}{|\sin(\pi\gamma)|}.$$

We also have the trivial bound $W_2 - W_1 + 1$. It follows, on writing $\xi := 2\beta m^2$, that

$$\sum_{\mu<w\le\nu} e(2\beta m^2 zw) \ll \min\left\{B, \frac{1}{|\sin(\pi\xi z)|}\right\},$$

and hence

$$|S_j(\alpha; B; m, a_j)|^2 \ll \sum_{|z|\le 2B} \min\left\{B, \frac{1}{|\sin(\pi\xi z)|}\right\}.$$

It is convenient to introduce at this point the periodic "saw-tooth" function

$$\theta\colon \mathbb{R} \to \left[-\tfrac{1}{2}, \tfrac{1}{2}\right], \quad \theta(t) := t - \max\left\{n \in \mathbb{Z}\colon n \le t + \tfrac{1}{2}\right\}.$$

This has the property that $|\theta(t)| \ll |\sin \pi t| \ll |\theta(t)|$. We proceed to decompose the available range for z into subsets

$$Z_h := \left\{z \in \mathbb{Z}\colon |z| \le 2B, \frac{h}{B} \le \theta(\xi z) < \frac{h+1}{B}\right\},$$

where h runs over integers with $|h| \le 1 + B/2$. Then if $|h| \ge 2$ we have

$$\frac{1}{|\sin(\pi\xi z)|} \ll \frac{1}{|\theta(\xi z)|} \ll \frac{B}{1+|h|}$$

for $z \in Z_h$, while if $|h| \le 1$ we have

$$B \ll \frac{B}{1+|h|}.$$

Thus in either case we find that

$$\min\left\{B, \frac{1}{|\sin(\pi\xi z)|}\right\} \ll \frac{B}{1+|h|},$$

whence

$$|S_j(\alpha; B; m, a_j)|^2 \ll (\max_h \#Z_h) \sum_{|h|\le 1+B/2} \frac{B}{1+|h|} \ll (\max_h \#Z_h) B \log B.$$

Now if $z_1, z_2 \in Z_h$ it follows that either $z_2 - z_1$ or $z_1 - z_2$ must be in Z_0, and hence that $\#Z_h \le 1 + 2\#Z_0$ for every value of h. Since $0 \in Z_0$ we have $1 + 2\#Z_0 \le 3\#Z_0$. We may therefore summarise our findings as follows.

LEMMA 3.2. *With the notation above we have*

$$|S_j(\alpha; B; m, a_j)|^2 \ll B(\log B)\#\{z \in \mathbb{Z}: |z| \le 2B, 0 \le \theta(\xi z) < B^{-1}\}.$$

Before proceeding further with our treatment of the minor arcs we pause to deduce an important corollary of this lemma. If we choose

$$\kappa_j = m, \quad \lambda_j = 2m, \quad \alpha = \frac{a}{q} \quad \text{and} \quad B = q$$

then we will have $\xi = 2m^2 c_j a/q$ and

$$S_q(a; m, \mathbf{a}) = S(\alpha; B\mathcal{R}; m, \mathbf{a}).$$

Moreover we find that

$$S_q(a; m, \mathbf{a}) = S(\alpha; B\mathcal{R}; m, \mathbf{a}) = \prod_{j=1}^{n} S_j(\alpha; B; m, a_j),$$

by (17). Lemma 3.2 shows that

$$|S_j(\alpha; B; m, a_j)|^2 \ll q(\log q)\#\{z \in \mathbb{Z}: |z| \le 2q, 0 \le \theta(2m^2 c_j az/q) < q^{-1}\}.$$

Since $2m^2 c_j az/q$ is a fraction with denominator q, the values of z we have to consider are precisely those for which $q \mid 2m^2 c_j az$. However a and q are coprime, so the condition reduces to $q \mid 2m^2 c_j z$. Moreover m and the various c_j are fixed, so that there are at $O(1)$ values of z in any range of length q. It follow that

$$|S_j(\alpha; B; m, a_j)|^2 \ll q(\log q)$$

and hence that

$$S_q(a; m, \mathbf{a}) \ll q^{n/2}(\log q)^{n/2}.$$

This therefore verifies the hypothesis made in our second lecture, as soon as $n \ge 5$, with any η in the range $0 < \eta < \frac{1}{2}$.

We must now complete our estimation of $S_j(\alpha; B; m, a_j)$ in the case in which α lies in the minor arc set $\mathfrak{m}$. We should first recall that the major arcs were the union of intervals

$$I(a, q) = \left[\frac{a}{q} - \frac{Q}{B^2}, \frac{a}{q} + \frac{Q}{B^2}\right], \quad (q \le Q),$$

where $Q = B^\delta$. We proceed to use Dirichlet's approximation theorem to find a rational approximation u/v to α satisfying

$$\left|\alpha - \frac{u}{v}\right| \le \frac{Q}{B^2 v}, \quad (u, v) = 1, \ v \le B^2 Q^{-1}. \tag{22}$$

It follows from this that $|\alpha - u/v| \leq QB^{-2}$, and since $\alpha \notin \mathfrak{M}$ we deduce that

$$v > Q. \tag{23}$$

Now for $|z| \leq 2B$ we have

$$\left|\xi z - \frac{2m^2 c_j uz}{v}\right| \leq \frac{4m^2|c_j|Q}{Bv} \leq \frac{4m^2|c_j|}{B},$$

by (23). Thus if $0 \leq \theta(\xi z) \leq B^{-1}$ we must have

$$\left|\theta\left(\frac{2m^2 c_j uz}{v}\right)\right| \leq \frac{1 + 4m^2|c_j|}{B}.$$

It follows that

$$2m^2 c_j uz \equiv l \,(\mathrm{mod}\, v) \tag{24}$$

for some integer l satisfying

$$|l| \leq v\frac{1 + 4m^2|c_j|}{B} \ll vB^{-1}.$$

The number of possible values for l is therefore $O(1 + vB^{-1})$. Moreover for each l the congruence (24) determines $O(1)$ residue class modulo v, since u and v are coprime and $2m^2c_j$ is a fixed non-zero integer. Finally the range $|z| \leq 2B$ contains at most $1 + 4Bv^{-1}$ integers from each residue class modulo v. On combining this information we find that

$$\begin{aligned}\#\{z \in \mathbb{Z} : |z| \leq 2B, 0 \leq \theta(\xi z) < B^{-1}\} &\ll (1 + vB^{-1})(1 + v^{-1}B) \\ &\ll 1 + vB^{-1} + v^{-1}B \ll BQ^{-1},\end{aligned}$$

in view of (22) and (23).

Finally we can apply Lemma 3.2 to deduce that

$$|S_j(\alpha; B; m, a_j)|^2 \ll B(\log B).BQ^{-1},$$

whence

$$|S_j(\alpha; B; m, a_j)| \ll B^{1-\delta/2}(\log B)^{1/2}. \tag{25}$$

On combining this with (18) and (20) we deduce, in the case $d = 2$, that

$$\left|\int_{\mathfrak{m}} S(\alpha; B\mathcal{R}; m, \mathbf{a})\, d\alpha\right| \ll \{B^{1-\delta/2}(\log B)^{1/2}\}^{n-4} B^{2+2\varepsilon}$$

for any fixed $\varepsilon > 0$. This gives us a satisfactory bound $o(B^{n-2})$ on choosing ε small enough, providing that $n \geq 5$. This completes the proof of Theorem 1.1.

4. Combining Analytic and Geometric Methods

We have seen in the previous lectures that when the circle method succeeds it actually proves that the rational points on the variety under consideration are dense in the adèlic points. However there are many varieties which do not have this latter property, because there is a non-empty "Brauer–Manin obstruction." Consequently the circle method is doomed to failure in such cases. While it would be out of place in these lectures to give a full description of the Brauer–Manin obstruction, it is hoped that the example below will convey an idea of the general situation.

In this lecture we shall examine two cases in which analytic methods have been successfully combined with a "descent" process to rescue the above situation. The first of these is due to Heath-Brown and Skorobogatov (Heath-Brown and Skorobogatov, 2002). Let K be an algebraic number field of degree $d \geq 2$, with integral basis $\omega_1, \dots, \omega_d$, and define a form of degree d by setting

$$N(\mathbf{x}) := N(x_1, \dots, x_d) := N_{K/\mathbb{Q}}(\omega_1 x_1 + \cdots + \omega_d x_d),$$

with the obvious abuse of notation. One is then interested in rational solutions to the equation

$$ct(t-1) = N(\mathbf{x}) \neq 0, \tag{26}$$

where c is a given non-zero integer, and where t and $\mathbf{x}$ are required to satisfy certain local conditions. (More generally, one can study $ct^a(t-1)^b = N(\mathbf{x}) \neq 0$ for arbitrary fixed integer exponents a and b.) These varieties provide a number of examples in which the rational points *fail* to be equidistributed.

Let us examine the case in which $K = \mathbb{Q}(\cos(2\pi/7))$. This is the real subfield of the cyclotomic field generated by the primitive 7th roots of unity. It is an Abelian cubic field. We shall use in particular the fact that $n \in \mathbb{N}$ is a norm from K if and only it is a product of prime powers p^e such that $p^e \equiv \pm 1 \pmod 7$ whenever $p \neq 7$. We shall take $c = 2$. Then there are solutions to (26) with $t = 28$, and also with $t = 29$. Thus there is a 2-adic solution (namely $t = 28$) with $|t|_2 < 1$, and a 7-adic solution (namely $t = 29$) with $|t-1|_7 < 1$. This produces an adèlic solution satisfying both constraints simultaneously.

We proceed to show however that there is no rational solution satisfying both constraints. To do this we write $t = u/v$ with u, v being coprime integers. We then see that $2u(u-v)v$ is a norm. It follows from the coprimality that exactly one of u, $u-v$ and v is even whence, using the coprimality condition again, we may conclude that one of the triples $u/4$, $u-v$, v, or u, $(u-v)/4$, v or u, $u-v$, $v/4$ consists entirely of integral norms. In each case this gives rise to an equation of the type $\pm 4N_1 \pm N_2 \pm N_3 = 0$, where each N_i is a norm. Since we must have $N_i \equiv 0$ or $\pm 1 \pmod 7$ we deduce that we must have $7|N_1$. It follows that whichever one of $u, u-v$ or v is even must also be a multiple of 7. Hence if $t = u/v$ satisfies $|t|_2 < 1$ we must also have $|t|_7 < 1$. Thus we cannot have $|t|_2 < 1$ and $|t - 1|_7 < 1$ simultaneously.

We have therefore shown that, for this example, the rational points are not dense in the adèlic points. Very roughly the argument consisted of using divisibility information to pass to a "descent variety" (which in this case was $\pm 4N_1 \pm N_2 \pm N_3 = 0$), for which there are local constraints not implicit in the original variety.

We proceed to examine the general situation. In homogeneous form the equation (26) becomes

$$cT(T - U)U^{d-2} = N(\mathbf{X}) \neq 0, \tag{27}$$

which involves $d + 2$ variables. In our preliminary discussion of the circle method we saw that one generally requires more that $2d$ variables for the method to succeed, so one cannot expect the circle method to work here. Indeed it transpires that there is in general a non-empty Brauer–Manin obstruction for this variety.

One way to think of the descent process for the variety (27) is to introduce a non-zero parameter λ, and to look for solutions in which

$$T = \lambda N(\mathbf{y}), \quad U = w^d$$

and

$$c\lambda^2 N(\mathbf{y}) - N(\mathbf{z}) = c\lambda w^d \tag{28}$$

for suitable non-zero integral $\mathbf{y}$, $\mathbf{z}$ and w. With the above choices we see that

$$cT(T - U)U^{d-2} = w^{d(d-2)} N(\mathbf{y})N(\mathbf{z}) \neq 0,$$

and since the set of norms is closed under multiplication this provides us with a solution to (27). It turns out that, if the local conditions on t and $\mathbf{x}$ are acceptable for the Brauer–Manin condition, then one can always find a value for λ for which (28) is locally solvable.

These manoeuvres have had two positive effects. Firstly they have replaced the original problem (27) with a new one (28), in which the number of variables, namely $2d + 1$, is now strictly greater than $2d$. Thus there is at least some prospect of the circle method working. The second point is a more subtle one. It turns out that there is no Brauer–Manin obstruction for varieties of the form (28), so again there is no *a priori* reason for the circle method to fail.

We shall not give the full argument needed to handle (28) by the circle method, but there is one point that does deserve to be discussed. It is far from true that the circle method works precisely when the number of variables satisfies $n > 2d$. In rare cases one can handle fewer variables; more frequently success requires distinctly more variables. Thus it is worthwhile seeing how, in this case, we can manage with as few as $2d + 1$ variables.

We saw, in the second lecture, how the major arcs could be handled under quite general conditions. The real difficulty lies with the minor arcs. For (28) it is natural to use a generating function

$$\sum_{\mathbf{y},\mathbf{z},w} e(\alpha\{c\lambda^2 N(\mathbf{y}) - N(\mathbf{z}) - c\lambda w^d\}),$$

with suitable summation conditions for the variables. This generating function will factor as

$$S_1(c\lambda^2\alpha)S_2(-\alpha)S_3(-c\lambda\alpha),$$

where

$$S_1(\beta) = \sum_{\mathbf{y}} e(\beta N(\mathbf{y})),$$
$$S_2(\beta) = \sum_{\mathbf{z}} e(\beta N(\mathbf{z}))$$

and

$$S_3(\beta) = \sum_{w} e(\beta w^d).$$

The minor arc procedure which led to (18) now shows that

$$\left|\int_{\mathfrak{m}} S_1(c\lambda^2\alpha)S_2(-\alpha)S_3(-c\lambda\alpha)\,d\alpha\right| \le \sup_{\alpha\in\mathfrak{m}} |S_3(-c\lambda\alpha)| \max_{k=1,2} \int_0^1 |S_k(c_k\alpha)|^2\,d\alpha,$$

where $c_1 = c\lambda^2$ and $c_2 = -1$.

The sum S_3 is the degree d version of the quadratic sum we investigated in (25), and may be handled by an extension of Weyl's method. It is the mean-value of S_k for $k = 1$ and 2, which interests us here. On expanding the square and integrating termwise we find, in complete analogy to (19), that

$$\int_0^1 |S_k(c_k\alpha)|^2\,d\alpha = \#\{\mathbf{y}_1,\mathbf{y}_2 \colon N(\mathbf{y}_1) = N(\mathbf{y}_2)\}$$

with the vectors $\mathbf{y}_1, \mathbf{y}_2$ in appropriate ranges. If we take, for illustrative purposes, all coordinates of $\mathbf{y}_1, \mathbf{y}_2$ to lie in the range $[1, B]$, we see that the trivial bound for $S_k(c_k\alpha)$ is $O(B^d)$. On the other hand, for each given $\mathbf{y}_1$ we claim that the number of $\mathbf{y}_2$ with $N(\mathbf{y}_1) = N(\mathbf{y}_2)$ is $O(B^\varepsilon)$, for any small fixed $\varepsilon > 0$. It follows that

$$\int_0^1 |S_k(c_k\alpha)|^2\,d\alpha \ll B^{d+\varepsilon},$$

so that we essentially have square-root cancellation. It is this remarkably sharp mean-value bound which is the key to success.

To verify the claim above we consider the equation $N_{K/\mathbb{Q}}(\gamma) = a$ with γ an algebraic integer in K, whose various conjugates are all $O(B)$. The number of integral ideals of norm a is at most $\tau(a)^d$, where $\tau(a)$ is the number of divisors of a. According to Lemma 3.1 this is $O(a^{d\varepsilon})$ for any $\varepsilon > 0$. Allowing for possible associates, each ideal of norm a corresponds to $O((\log B)^d)$ possible integers γ, in view of the constraints on the conjugates of γ. The claimed result then follows, on re-defining ε.

The remainder of this lecture is devoted to a second example in which analytic methods must be combined with a "descent" argument. Full details are in the author's paper (Heath-Brown, 2003).

In this case we shall examine a variety defined by a pair of simultaneous equations of the shape

$$V\colon \begin{cases} L_1(x_1,x_2)L_2(x_1,x_2) = x_3^2 + x_4^2 \\ L_3(x_1,x_2)L_4(x_1,x_2) = x_5^2 + x_6^2 \end{cases} \tag{29}$$

where $L_1, \ldots, L_4$ are linear forms defined over $\mathbb{Z}$, no two of which are proportional. It is known that the variety V can fail to satisfy the Hasse Principle, and that even if there are rational points, they may fail to be dense in the adèlic points. Thus a direct application of the circle method is impossible. Indeed for pairs of quadratic forms the best result in the literature is that of (Cook, 1971), which relates only to diagonal forms, and requires at least 9 variables.

We shall consider points on V (more precisely, on the affine cone over V), for which $\mathbf{x} = (x_1, x_2) \in B\mathcal{R}$. Here $\mathcal{R} \subset \mathbb{R}^2$ is open, bounded and convex, with a piecewise continuously differentiable boundary, and B is a large positive parameter. Note that if $\mathbf{x} \in B\mathcal{R}$ then the sizes of $x_3, \ldots, x_6$ are all restricted to be $O(B)$. We further assume that $L_i(\mathbf{x}) > 0$ for $1 \le i \le 4$, for all $\mathbf{x} \in \mathcal{R}$. One further technical condition needs to be imposed, namely that

$$L_1(x_1,x_2) \equiv L_2(x_1,x_2) \equiv \nu x_1 \pmod 4$$

and

$$L_3(x_1,x_2) \equiv L_4(x_1,x_2) \equiv \nu' x_1 \pmod 4,$$

for appropriate $\nu, \nu' = \pm 1$. It is then natural to impose a congruence restriction on $\mathbf{x}$ and to work with

$$(B\mathcal{R})^{(2)} := \{\mathbf{x} \in B\mathcal{R}\colon x_1 \equiv 1 \pmod 2\}.$$

Introducing the arithmetic function

$$r(n) := \#\{(a,b) \in \mathbb{Z}^2\colon a^2 + b^2 = n\},$$

we see that the number of solutions of (29) under consideration is given by

$$S(B) := \sum_{\mathbf{x}\in(B\mathcal{R})^{(2)}} r(L_1(\mathbf{x})L_2(\mathbf{x}))r(L_3(\mathbf{x})L_4(\mathbf{x})).$$

The principle result of (Heath-Brown, 2003) is then the following.

THEOREM 4.1. *Let σ_∞ and σ_p be the local densities for the variety V with equations* (29), *for the set* $(B\mathcal{R})^{(2)}$. *If $\sigma_p = 0$ for any prime p, then V has no rational point with $(x_1, x_2) \in (B\mathcal{R})^{(2)}$. If $\sigma_p \neq 0$ for every prime p, then*

$$S(B) = \{1+\varepsilon\}\sigma_\infty \prod_p \sigma_p + o(B^2)$$

for a certain rational constant $\varepsilon \in [-1, 1]$.

If $\varepsilon = -1$ then V has no rational point with $(x_1, x_2) \in (B\mathcal{R})^{(2)}$.

Note that σ_∞ will be a positive constant multiple of B^2, so that we have a genuine asymptotic formula if the other local factors are positive, and if $\varepsilon \neq -1$. In fact the constant ε is given explicitly. It is a finite product of certain local factors. We therefore see how the Hardy–Littlewood asymptotic formula holds with the additional factor $1+\varepsilon$, which reflects Brauer–Manin considerations. Indeed, if all the σ_p are non-zero, one has $1+\varepsilon = 0$ precisely when there is a Brauer–Manin obstruction.

Just as we saw how (27) led to (28), we may replace the variety (29) by the system

$$\begin{aligned} L_1(\mathbf{x}) &= d(y_1^2+y_2^2), \quad & L_2(\mathbf{x}) &= d(y_3^2+y_4^2), \\ L_3(\mathbf{x}) &= d'(y_5^2+y_6^2), \quad & L_4(\mathbf{x}) &= d'(y_7^2+y_8^2), \end{aligned} \tag{30}$$

for different choices $d, d' \in \mathbb{N}$. Here we use the fact that $d(y_1^2+y_2^2)\times d(y_3^2+y_4^2)$ is a sum of two squares, $x_3^2+x_4^2$, say. The underlying process can be expressed more precisely using the identity

$$r(mn) = \frac{1}{4}\sum_{d|m,n} \mu(d)\chi(d)r(m/d)r(n/d),$$

where χ is the non-principal character modulo 4. This shows that

$$S(B) = \frac{1}{16}\sum_{d,d'} \mu(d)\mu(d')\chi(dd')S(B; d, d'), \tag{31}$$

with

$$S(B; d, d') := \sum_{\mathbf{x}\in(B\mathcal{R})^{(2)}} r(L_1(\mathbf{x})/d)r(L_2(\mathbf{x})/d)r(L_3(\mathbf{x})/d')r(L_4(\mathbf{x})/d').$$

Here we set $r(q) = 0$ if q is not an integer. The inner sum on the right clearly counts solutions to (30).

Just to check on progress, let us suppose one were to eliminate $\mathbf{x} = (x_1, x_2)$ from the equations (30). This would produce a pair of diagonal quadratic equations in the 8 variables $y_1, \ldots, y_8$. Now we have noted already that Cook's work (Cook, 1971) would allow us to handle such a system via the circle method, if only there were 9 or more variables. However it transpires that although the circle method is not itself available for the varieties (30) there is an alternative analytic method which one can use, based on the fact that

$$r(m) = 4\sum_{d|m}\chi(d).$$

The argument is quite delicate, and produces a result of the form

$$S(B; d, d') = C_{d,d'}B^2 + O(B^2(\log B)^{-\delta}) \tag{32}$$

for some small positive constant δ, for any fixed $d, d' \in \mathbb{N}$. It should be observed that our proof of Theorem 1.1 saves a positive power of the parameter B, while in contrast the above estimate saves only a power of $\log B$. This is a reflection of the fact that we are working very much on the border of what is currently feasible. Although the result does not come from the circle method, the constant $C_{d,d'}$ is none the less the product of the local densities for the equations (30). Thus we have exactly the result one would expect from the circle method, even though the method of proof is different.

Once (32) has been established it is possible to sum up the various terms in (31), using the asymptotic formula (32) for each pair d, d'. (There are important issues concerning uniformity in d ,d' which need to be dealt with, but these can in fact be overcome.) This leads to a result of the form

$$\sum_{\mathbf{x}\in(B\mathcal{R})^{(2)}} r(L_1(\mathbf{x})L_2(\mathbf{x}))r(L_3(\mathbf{x})L_4(\mathbf{x})) = CB^2 + o(B^2)$$

with

$$C = \frac{1}{16}\sum_{d,d'}\mu(d)\mu(d')\chi(dd')C_{d,d'}.$$

To complete the proof one then computes this sum, and relates it to the product of local densities for (29).

We conclude with a numerical illustration, relating to the system

$$x_1(x_1 + 12x_2) = x_3^2 + x_4^2, \quad (x_1 + 4x_2)(x_1 + 16x_2) = x_5^2 + x_6^2.$$

In this case one can show that $0 < 1 + \varepsilon < 2$ and that

$$\{1 + \varepsilon\}\sigma_\infty \prod_p \sigma_p = 2$$

for the region

$$\mathcal{R} = \{0 < x_1, x_1 + 16x_2 < 1\}.$$

(It would appear that the occurrence of an integer value for the overall density is no more than a fluke!) One now has the numerical values of Table I.

TABLE I.

B	$S(B)$	$S(B)/2B^2$
1000	1993472	0.9967...
2000	8030592	1.0038...
4000	32057728	1.0018...
8000	1276046726	0.9969...
16000	511437824	0.9989...
32000	2043518720	0.9978...

Acknowledgements

Part of this work was undertaken while the author was attending the Diophantine Geometry session at the Centro di Ricerca Matematica Ennio De Giorgi in Pisa. The hospitality and financial support of the centre is gratefully acknowledged.

Thanks are also due to Dr. T. Browning, whose careful reading of a first draft of these notes led to many improvements and corrections.

References

Birch, B. J. (1961/1962) Forms in many variables, *Proc. Roy. Soc. Ser. A* **265**, 245–263.

Cook, R. J. (1971) Simultaneous quadratic equations, *J. London Math. Soc. (2)* **4**, 319–326.

Davenport, H. (2005) Analytic methods for Diophantine equations and Diophantine inequalities, In *Cambridge Mathematical Library*, Cambridge, Cambridge University Press.

Davenport, H. and Lewis, D. J. (1969) Simultaneous equations of additive type, *Philos. Trans. Roy. Soc. London Ser. A* **264**, 557–595.

Deligne, P. (1980) La conjecture de Weil. I, *Inst. Hautes Études Sci. Publ. Math.* **52**, 137–252.

Ford, K. (1995) New estimates for mean values of Weyl sums, *Internat. Math. Res. Notices* pp. 155–171.

Heath-Brown, D. R. (1996) A new form of the circle method, and its application to quadratic forms, *J. Reine Angew. Math.* **481**, 149–206.

Heath-Brown, D. R. (2003) Linear relations amongst sums of two squares, In *London Math. Soc. Lecture Note Ser.*, Vol. 303 of *London Math. Soc. Lecture Note Ser.*, Cambridge, pp. 133–176, Cambridge Univ. Press.

Heath-Brown, D. R. and Skorobogatov, A. N. (2002) Rational solutions of certain equations involving norms, *Acta Math.* **189**, 161–177.

Hooley, C. (1994) On nonary cubic forms. III, *J. Reine Angew. Math.* **456**, 53–63.

Vaughan, R. C. (1986) On Waring's problem for cubes, *J. Reine Angew. Math* **365**, 122–170.

Vaughan, R. C. (1997) The Hardy–Littlewood method, In *Cambridge Tracts in Mathematics*, Vol. 125, Cambridge-New York, Cambridge University Press.

Wang, Y. (1991) *Diophantine equations and inequalities in algebraic number fields*, Berlin, Springer-Verlag.

Weyl, H. (1916) Über die Gleichverteilung von Zahlen mod Eins, *Math. Ann.* **77**, 313–352.

UNIVERSAL TORSORS OVER DEL PEZZO SURFACES AND RATIONAL POINTS

Ulrich Derenthal and Yuri Tschinkel
Universität Göttingen

Abstract. We discuss Manin's conjecture (with Peyre's refinement) concerning the distribution of rational points of bounded height on Del Pezzo surfaces, by highlighting the use of universal torsors in such counting problems. To illustrate the method, we provide a proof of Manin's conjecture for the unique split singular quartic Del Pezzo surface with a singularity of type $\mathbf{D}_4$.

1. Introduction

Let $f \in \mathbb{Z}[x_0, \dots, x_n]$ be a non-singular form of degree d. By the circle method,

$$N(f, B) := \#\left\{\mathbf{x} \in \mathbb{Z}^{n+1}/\pm \mid \max_j(|x_j|) \leqslant B\right\} \sim c \cdot B^{n+1-d}$$

(where $\mathbf{x} \in \mathbb{Z}^{n+1}/\pm$ means that we identify $\mathbf{x}$ with $-\mathbf{x} = (-x_0, \dots, -x_n)$) with $c \in \mathbb{R}_{>0}$, provided that $n \geqslant 2^d \cdot (d-1)$, and $f(\mathbf{x}) = 0$ has solutions over all completions of $\mathbb{Q}$ (see (Birch, 1962)). Let $X = X_f \subset \mathbb{P}^n$ be the smooth hypersurface over $\mathbb{Q}$, given by $f(\mathbf{x}) = 0$. It follows that

$$N(X, -K_X, B) \; = \#\{\mathbf{x} \in X(\mathbb{Q}) \mid H_{-K_X}(\mathbf{x}) \leqslant B\} \sim C \cdot B, \tag{1}$$

as $B \to \infty$. Here $X(\mathbb{Q})$ is the set of rational points on X, represented by primitive vectors $\mathbf{x} \in (\mathbb{Z}^{n+1}_{\text{prim}} \setminus 0)/\pm$ (i.e., $\mathbf{x} = (x_0, \dots, x_n)$ is identified with $-\mathbf{x}$, and there is no prime dividing all coordinates $x_0, \dots, x_n$), and

$$H_{-K_X}(\mathbf{x}) := \max_j(|x_j|)^{n+1-d}, \quad \text{for } \mathbf{x} = (x_0, \dots, x_n) \in (\mathbb{Z}^{n+1}_{\text{prim}} \setminus 0)/\pm\,. \tag{2}$$

is the *anticanonical height* of a primitive representative.

In 1989 Manin initiated a program towards understanding connections between certain geometric invariants of algebraic varieties over number fields and their arithmetic properties, in particular, the distribution of rational points of bounded height, see (Franke et al., 1989) and (Batyrev and Manin, 1990).

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 169–196.

The main goal is an extension of the asymptotic formula (1) to other algebraic varieties of *small* degree, called Fano varieties, which are not necessarily isomorphic to hypersurfaces in projective space.

It became apparent that, in general, to obtain a geometric interpretation of asymptotic results, it may be necessary to restrict to appropriate Zariski open subsets of X. Otherwise, the number of rational points on a Zariski closed subset of lower dimension may dominate the total number of rational points; e.g., this phenomenon occurs for the surface (4) below where there are many, in fact too many, rational points on lines on this surface. These are easy to count, so to make things more interesting, we count the number of rational points on the complement of these lines. It is often interesting to count rational points in finite extensions of the rationals: while $X(\mathbb{Q})$ might be empty, $X(k)$ could still contain infinitely many points for some number field k.

Of particular interest are Del Pezzo surfaces (cf. (Manin, 1986)), e.g., cubic surfaces $S_3 \subset \mathbb{P}^3$ or degree 4 surfaces $S_4 := Q_1 \cap Q_2 \subset \mathbb{P}^4$, where Q_1, Q_2 are *quadrics* (defined by homogeneous equations of degree 2 in $x_0, \dots, x_4$). Geometrically, smooth Del Pezzo surfaces are obtained by blowing up $\leqslant$ 8 *general points*[1] in $\mathbb{P}^2$. Blowing up is a standard procedure in algebraic geometry (cf. (Hartshorne, 1977, Section I.4)). The blow-up $\pi\colon S' \to S$ of a surface S at a point p replaces p by a curve E in a particular way. We have $S \setminus \{p\} \cong S' \setminus E$, so S and S' are *birationally equivalent*. In our situation, this shows that Del Pezzo surfaces are birational to $\mathbb{P}^2$, provided the ground field is algebraically closed.

We can think of *divisors* on blow-ups S of $\mathbb{P}^2$ as formal sums of curves on S. Considering divisors up to a certain equivalence relation (see (Hartshorne, 1977, Section II.6)) leads to the *Picard group* $\mathrm{Pic}(S)$ of divisor classes on S.

For two curves on S which intersect transversally,[2] their *intersection number* is the number of intersection points. As explained in (Hartshorne, 1977, Section V.1), this can be extended to arbitrary divisor classes (by defining the non-degenerate *intersection form* $(\cdot,\cdot)$ on $\mathrm{Pic}(S)$). In particular, this defines the *self intersection number* (E, E) of (the class of) a curve E. Of special interest are irreducible curves for which this number is negative, which are called *exceptional curves*. Arithmetically, rational points tend to accumulate on exceptional curves where they are easy to count. The main focus of this paper is to count rational points on the complement of the exceptional curves.

[1] no three points on a line, no six points on a curve of degree 2, no eight points with one of them singular on a curve of degree 3

[2] I.e., all intersections have multiplicity one.

For smooth Del Pezzo surface of degree 3 and 4, the exceptional curves are exactly the lines (in the standard embedding considered above), having self intersection number -1.

The singular Del Pezzo surfaces are obtained as follows: we blow up $\mathbb{P}^2$ in special configurations of points (e.g., three points on a line). This results in a smooth surface $\widetilde{S}$ containing exceptional curves with self intersection number -2 (called (-2)-curves; we do not subsequently blow up points on (-2)-curves). Contracting the (-2)-curves (which is the opposite of blowing up) gives a singular Del Pezzo surface S whose *minimal desingularization* is $\widetilde{S}$. For the surface (4) below, more details can be found in Section 2.

For number fields, we say that a Del Pezzo surface is *split* if all of the exceptional curves are defined over that ground field, in which case the surface is birational to $\mathbb{P}^2$. There do exist *non-split* Del Pezzo surfaces which are birational to $\mathbb{P}^2$ over that ground field; however, the generic Del Pezzo surface is non-split and is not birational to $\mathbb{P}^2$ over the ground field.

From now on, we work over $\mathbb{Q}$. Manin's conjecture in the special case of Del Pezzo surfaces can be formulated as follows.

CONJECTURE 1.1. *Let S be a Del Pezzo surface with at most rational double points*[3] *over $\mathbb{Q}$. Then there exists a subset $S^0 \subset S$ which is dense and open in the Zariski topology such that*

$$N(S^0, -K_S, B) \sim c_{S,H} \cdot B(\log B)^{r-1}, \tag{3}$$

as $B \to \infty$, where r is the rank of the Picard group of the minimal desingularization $\widetilde{S}$ of S, over $\mathbb{Q}$.

The constant $c_{S,H}$ has been defined by Peyre (Peyre, 1995); it should be non-zero if $S(\mathbb{Q}) \neq \emptyset$. It is analogous to the singular series and the singular integral that you meet in the classical circle method (see Heath-Brown's article in this volume (Heath-Brown, 2006)). Note that a line defined over $\mathbb{Q}$ on a Del Pezzo surface such as S_3 or S_4 contributes $\sim c \cdot B^2$ rational points to the counting function (for some positive constant c). Thus it is expected that S^0 is the complement to all lines defined over $\mathbb{Q}$ (i.e., the exceptional curves).

Table I gives an overview of current results towards Conjecture 1.1 for Del Pezzo surfaces. In Column 4 ("type of result"), "asymptotic" means that the analog of (3) is established, including the predicted value of the constant; "bounds" means that only upper and lower bounds of the expected order of magnitude $B(\log B)^{r-1}$ are known.

[3] These are "mild" singularities which can be dealt with. The technique is to "resolve" these singularities by replacing each of them by a curve whose irreducible components are isomorphic to $\mathbb{P}^1$.

TABLE I. Results for Del Pezzo surfaces

degree	singularities	(non-)split	type of result	reference
$\geqslant 6$	–	split	asymptotic	(Batyrev and Tschinkel, 1998)
5	–	split	asymptotic	(de la Bretèche, 2002)
5	–	non-split	asymptotic	(de la Bretèche and Fouvry, 2004)
4	$\mathbf{D}_5$	split	asymptotic	(Chambert-Loir and Tschinkel, 2002), (de la Bretèche and Browning, 2006)
4	$\mathbf{D}_4$	non-split	asymptotic	(de la Bretèche and Browning, 2005)
4	$\mathbf{D}_4$	split	asymptotic	this paper
4	$3\mathbf{A}_1$	split	bounds	(Browning, 2005)
3	$3\mathbf{A}_2$	split	asymptotic	(Batyrev and Tschinkel, 1998), (de la Bretèche, 1998), ...
3	$4\mathbf{A}_1$	split	bounds	(Heath-Brown, 2003)
3	$\mathbf{D}_4$	split	bounds	(Browning, 2004)
3	$\mathbf{E}_6$	split	asymptotic	(Derenthal, 2005), (de la Bretèche et al., 2005)

The paper (Batyrev and Tschinkel, 1998) contains a proof of Manin's conjecture for toric Fano varieties, including all smooth Del Pezzo surfaces of degree $\geqslant 6$ and the $3\mathbf{A}_2$ cubic surface[4]. This result also covers:

- all singular surfaces of degree $\geqslant 7$ (i.e., $\mathbf{A}_1$ in degree 7 and 8),
- $\mathbf{A}_1$, $2\mathbf{A}_1$, $\mathbf{A}_2 + \mathbf{A}_1$ in degree 6,
- $2\mathbf{A}_1$, $\mathbf{A}_2 + \mathbf{A}_1$ in degree 5,
- $4\mathbf{A}_1$, $\mathbf{A}_2 + 2\mathbf{A}_1$, $\mathbf{A}_3 + 2\mathbf{A}_1$ in degree 4.

Figure 1 shows all points of height $\leqslant 50$ on the Cayley cubic surface (Example 7.3), which has four singularities of type $\mathbf{A}_1$ and was considered in (Heath-Brown, 2003).

In Figure 2, we see points of height $\leqslant 1000$ on the $\mathbf{E}_6$ cubic surface ((Derenthal, 2005) and (de la Bretèche et al., 2005)).

[4] Our study of rational points on Del Pezzo surfaces involves classification of their singularities. In algebraic geometry, it is a basic result that singularities on Del Pezzo surfaces are labeled by Dynkin diagrams; the corresponding Dynkin diagram describes the number and intersection behaviour of the (-2)-curves on $\widetilde{S}$. For further explanation of the notations $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{D}_4, \mathbf{D}_5, \mathbf{E}_6$, we refer the interested reader to (Coray and Tsfasman, 1988).

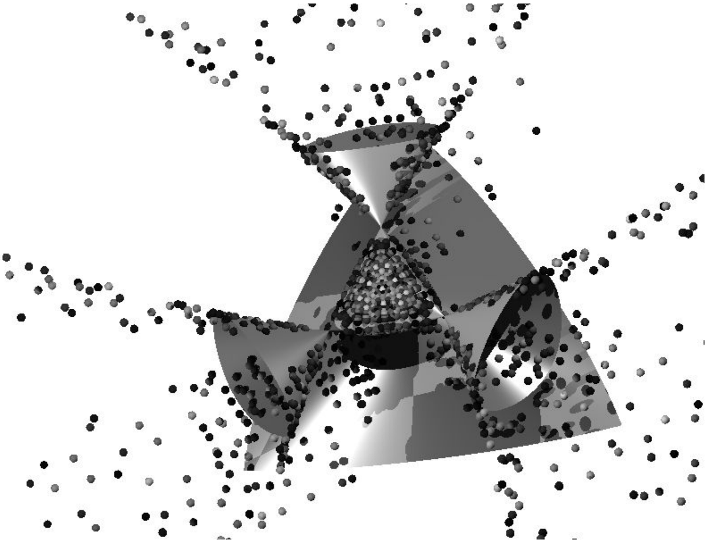


Figure 1. Points of height $\leqslant 50$ on the Cayley cubic surface $x_0x_1x_2 + x_0x_1x_3 + x_0x_2x_3 + x_1x_2x_3 = 0$.

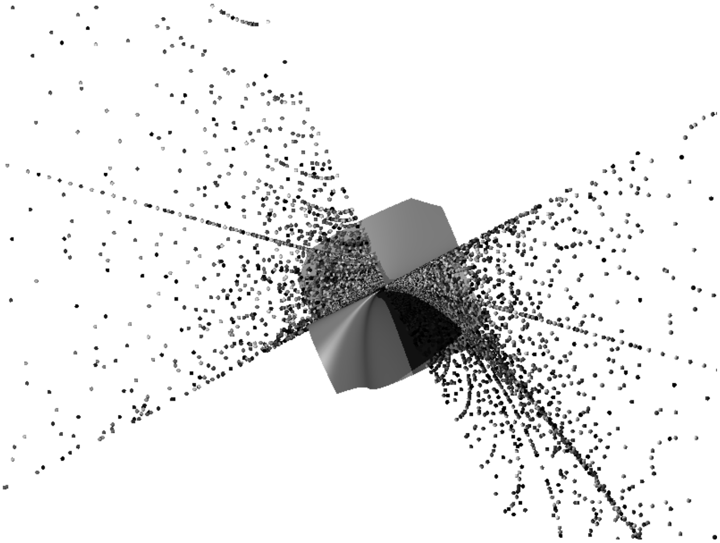


Figure 2. Points of height $\leqslant 1000$ on the $\mathbf{E}_6$ singular cubic surface $x_1x_2^2 + x_2x_0^2 + x_3^3 = 0$ with $x_0, x_2 > 0$.

The proofs of Manin's conjecture proceed either via the height zeta function

$$Z(s) := \sum_{\mathbf{x}\in X^\circ(\mathbb{Q})} H_{-K_X}(\mathbf{x})^{-s},$$

whose analytic properties are related to the asymptotic (3) by Tauberian theorems, or via the lifting of the counting problem to the *universal torsor*—an auxiliary variety parametrizing rational points. Experience shows that counting points on the universal torsor is often easier. The torsor approach has been developed by Colliot-Thélène and Sansuc in the context of the Brauer-Manin obstruction (Colliot-Thélène and Sansuc, 1987) and applied to Manin's conjecture by Peyre (Peyre, 1998) and Salberger (Salberger, 1998).

In the simplest case of hypersurfaces $X = X_f \subset \mathbb{P}^n$ over $\mathbb{Q}$, with $n \geqslant 4$, this is exactly the passage from rational vectors $\mathbf{x} = (x_0, \dots, x_n)$, modulo the diagonal action of $\mathbb{Q}^*$, to primitive lattice points $(\mathbb{Z}^{n+1}_{\text{prim}} \setminus 0)/\pm$. Geometrically, we have

$$\mathbb{A}^{n+1} \setminus 0 \xrightarrow{\mathbb{G}_{\mathrm{m}}} \mathbb{P}^n \quad \text{and} \quad \mathcal{T}_X \xrightarrow{\mathbb{G}_{\mathrm{m}}} X.$$

Here, $\mathcal{T}_X$ is the hypersurface in $\mathbb{A}^{n+1} \setminus 0$ defined by the form f, the 1-dimensional torus $\mathbb{G}_{\mathrm{m}}$ is interpreted as the Néron-Severi torus T_{NS}[5] Rational points on the base surface X are lifted to integral points on the torsor, modulo the action of the group of units $T_{\mathrm{NS}}(\mathbb{Z}) = \{\pm 1\}$. The height inequality $H(\mathbf{x}) \leqslant B$ for $\mathbf{x}$ on the base X translates into the usual height inequality on the torsor (2). In this case, it is possible to count the points on the torsor using the classical circle method.

In general, a torsor under an algebraic torus T is determined by a homomorpism $\chi\colon \mathfrak{X}^*(T) \to \mathrm{Pic}(X)$ to the Picard group of the underlying variety X; the term *universal* is applied when χ is an isomorphism.

For hypersurfaces in $\mathbb{P}^3$, or more generally for complete intersection surfaces (i.e., where S is the intersection of k hypersurfaces in $\mathbb{P}^{k+2}$), the Picard group may have higher rank. For example, for split smooth cubic surfaces $S = S_3 \subset \mathbb{P}^3$ the rank is 7, so that the dimension of the corresponding universal torsor $\mathcal{T}_S$ is 9; for quartic Del Pezzo surfaces these are 6 and 8, respectively. Therefore, the counting problem is now reduced to a higher-dimensional variety.

It is expected that the passage to universal torsors, which can be considered as natural *descent varieties*, will facilitate the proof of Manin's conjecture (Conjecture 1.1), at least for Del Pezzo surfaces. Rational points on S are lifted to certain integral points on $\mathcal{T}_S$, modulo the action of $T_{\mathrm{NS}}(\mathbb{Z}) =$

[5] An algebraic torus whose characters $\mathfrak{X}^*(T_{\mathrm{NS}})$ are isomorphic to the Picard group (lattice) of $\mathbb{P}^n$, resp. X, and the map is the natural quotient by its (diagonal) action.

$(\pm 1)^r$, where r is the rank of $\mathrm{Pic}(S)$, and the height inequality on S translates into appropriate inequalities on $\mathcal{T}_S$. This explains the interest in the projective geometry of torsors, and expecially, in their equations. The explicit determination of these equations is an interesting algebro-geometric problem, involving tools from invariant theory and toric geometry.

In this note, we illustrate the torsor approach to asymptotics of rational points in the case of a particular singular surface $S \subset \mathbb{P}^4$ of degree 4 given by:

$$x_0x_3 - x_1x_4 = x_0x_1 + x_1x_3 + x_2^2 = 0. \tag{4}$$

This is a split Del Pezzo surface, with a singularity of type $\mathbf{D}_4$.

THEOREM 1.2. *The number of $\mathbb{Q}$-rational points of anticanonical height bounded by B on the complement S^0 of the $\mathbb{Q}$-rational lines on S (as defined in* (4)) *satisfies*

$$N(S^0, -K_S, B) = c_{S,H} \cdot B \cdot Q(\log B) + O(B(\log B)^3) \quad \text{as } B \to \infty,$$

where Q is a monic polynomial of degree 5, *and*

$$c_{S,H} = \frac{1}{34560} \cdot \omega_\infty \cdot \prod_p (1 - 1/p)^6(1 + 6/p + 1/p^2)$$

*is the constant predicted by Peyre (*Peyre, 1995*), with p running through all primes and*

$$\omega_\infty = 3 \iiint_{\{(t,u,v)\in\mathbb{R}^3 | 0\leqslant v\leqslant 1, |tv^2|, |v^2u|, |v(tv+u^2)|, |t(tv+u^2)|\leqslant 1\}} 1 \, \mathrm{d}t \, \mathrm{d}u \, \mathrm{d}v.$$

In (de la Bretèche and Browning, 2005), Manin's conjecture is proved for a non-split surface with a singularity of the same type. However, these results do not follow from each other.

In Section 2, we collect some facts about the geometric structure of S. In Section 3, we calculate the expected value of $c_{S,H}$ and show that Theorem 1.2 agrees with Manin's conjecture.

In our case, the universal torsor is an affine hypersurface. In Section 4, we calculate its equation, stressing the relation with the geometry of S. We make explicit the coprimality and the height conditions. The method is more systematic than the derivation of torsor equations in (de la Bretèche and Browning, 2006) and (de la Bretèche et al., 2005), and should bootstrap to more complicated cases, e.g., other split Del Pezzo surfaces.

Note that our method gives coprimality conditions which are different from the ones in (de la Bretèche and Browning, 2006) and (de la Bretèche et al., 2005), but which are in a certain sense more natural: they are related to the set of points on $\mathcal{T}_S$ which are *stable* with respect to the action of the Néron-Severi torus (in the sense of geometric invariant theory, cf., (Dolgachev, 2003) and (Hu and Keel, 2000)). Our conditions involve only coprimality of certain pairs of variables, while the other method produces a mix of square-free variables and coprimalities.

In Section 5, we estimate the number of integral points on the universal torsor by iterating summations over the torsor variables and using results of elementary analytic number theory. Finally we arrive at Lemma 5.3, which is very similar in appearance to (de la Bretèche and Browning, 2006, Lemma 10) and (Derenthal, 2005, Lemma 12). In Section 6 we use familiar methods of height zeta functions to derive the exact asymptotic. We isolate the expected constant $c_{S,H}$ and finish the proof of Theorem 1.2. In Section 7 we write down examples of universal torsors for other Del Pezzo surfaces and discuss their geometry.

2. Geometric Background

In this section, we collect some geometric facts concerning the surface S. We show that Manin's conjecture for S is not a special case of available more general results for Del Pezzo surfaces.

LEMMA 2.1. *The surface S has the following properties*:

1. *It has exactly one singularity of type* $\mathbf{D}_4$ *at* $q = (0 : 0 : 0 : 0 : 1)$.
2. *S contains exactly two lines*:

$$E_5 = \{x_0 = x_1 = x_2 = 0\} \quad \textit{and} \quad E_6 = \{x_1 = x_2 = x_3 = 0\},$$

which intersect in q.

3. *The projection from the line E_5 is a birational map*

$$\begin{aligned} \phi\colon S &\to \mathbb{P}^2 \\ \mathbf{x} &\mapsto (x_0 : x_2 : x_1) \end{aligned}$$

which is defined outside E_5. It restricts to an isomorphism between

$$S^0 = S \setminus (E_5 \cup E_6) = \{\mathbf{x} \in S \mid x_1 \neq 0\} \ \textit{and} \ \mathbb{A}^2 \cong \{(t : u : v) \mid v \neq 0\} \subset \mathbb{P}^2,$$

whose inverse is the restriction of

$$\begin{aligned} \psi\colon \mathbb{P}^2 &\to S, \\ (t : u : v) &\mapsto (tv^2 : v^3 : v^2u : -v(tv + u^2) : -t(tv + u^2)) \end{aligned}$$

Similar results hold for the projection from E_6.

4. *The process of resolving the singularity q gives four exceptional curves $E_1, \dots, E_4$ and produces the minimal desingularization $\widetilde{S}$, which is also the blow-up of $\mathbb{P}^2$ in five points.*
   Proof. Direct computations.

It will be important to know the details of the sequence of five blow-ups of $\mathbb{P}^2$ giving $\widetilde{S}$ as in Lemma 2.1(4):

In order to describe the points in $\mathbb{P}^2$, we need the lines

$$E_3 = \{v = 0\}, \quad A_1 = \{u = 0\}, \quad A_2 = \{t = 0\}$$

and the curve $A_3 = \{tv + u^2 = 0\}$.

LEMMA 2.2. *The following five blow-ups of $\mathbb{P}^2$ result in $\widetilde{S}$:*

- *Blow up the intersection of E_3, A_1, A_3, giving E_2.*
- *Blow up the intersection of E_2, E_3, A_3, giving E_1.*
- *Blow up the intersection of E_1 and A_3, giving E_4.*
- *Blow up the intersection of E_4 and A_3, giving E_6.*
- *Blow up the intersection of E_3 and A_2, giving E_5.*

Here, the order of the first four blow-ups is fixed, and the fifth blow-up can be done at any time.

The Dynkin diagram in Figure 3 *describes the final configuration of divisors $E_1, \dots, E_6$, A_1, A_2, A_3. Here, A_1, A_2, A_3 intersect at one point.*

The quartic Del Pezzo surface with a singularity of type $\mathbf{D}_4$ is not toric, and Manin's conjecture does not follow from the results of (Batyrev and Tschinkel, 1998). The $\mathbf{D}_5$ example of (de la Bretèche and Browning, 2006) is an equivariant compactification of $\mathbb{G}_a^2$ (i.e., S has a Zariski open subset isomorphic to $\mathbb{A}^2$, and the obvious action of $\mathbb{G}_a^2$ on this open subset extends to S), and thus a special case of (Chambert-Loir and Tschinkel, 2002).

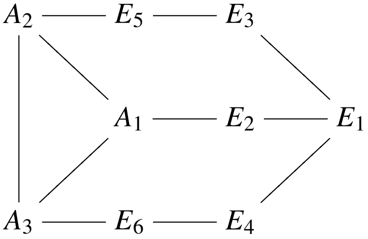


Figure 3. Extended Dynkin diagram.

LEMMA 2.3. *The quartic Del Pezzo surface with a singularity of type* $\mathbf{D}_4$ *is a compactification of* $\mathbb{A}^2$*, but not an equivariant compactification of* $\mathbb{G}_a^2$.

Proof. We follow the strategy of (Hassett and Tschinkel, 2004, Remark 3.3).

Consider the maps ϕ, ψ as in Lemma 2.1(3). As ψ restricts to an isomorphism between $\mathbb{A}^2$ and the open set $S^0 \subset S$, the surface S is a compactification of $\mathbb{A}^2$.

If S were an equivariant compactification of $\mathbb{G}_a^2$ then the projection ϕ from E_5 would be a $\mathbb{G}_a^2$-equivariant map, giving a $\mathbb{G}_a^2$-action on $\mathbb{P}^2$. The line $\{v = 0\}$ would be invariant under this action. The only such action is the standard translation action

$$\begin{aligned} \tau\colon \mathbb{P}^2 &\to \mathbb{P}^2, \\ (t : u : v) &\mapsto (t + \alpha v : u + \beta v : v). \end{aligned}$$

However, this action does not leave the linear series

$$(tv^2 : v^3 : v^2 u : -v(tv + u^2) : -t(tv + u^2))$$

invariant, which can be seen after calculating

$$\begin{aligned} t(tv + u^2) \ \mapsto \ & (t + \alpha v)((t + \alpha v)v + (u + \beta v)^2) \\ &= t(tv + u^2) + 2\beta tuv + (\beta^2 + \alpha)tv^2 + \alpha v(tv + u^2) \\ &\qquad + 2\alpha\beta v^2 u + (\alpha\beta^2 + \alpha^2)v^3, \end{aligned}$$

since the term tuv does not appear in the original linear series.

3. Manin's Conjecture

LEMMA 3.1. *Let S be the surface* (4). *Manin's conjecture for S states that the number of rational points of height $\leqslant B$ outside the two lines is given by*

$$N(S^0, -K_S, B) \sim c_{S,H} \cdot B(\log B)^5,$$

where $c_{S,H} = \alpha(S) \cdot \beta(S) \cdot \omega_H(S)$ *with*

$$\begin{aligned} \alpha(S) &= (5! \cdot 4 \cdot 2 \cdot 3 \cdot 3 \cdot 2 \cdot 2)^{-1} = (34560)^{-1} \\ \beta(S) &= 1 \\ \omega_H(S) &= \omega_\infty \cdot \prod_p (1 - 1/p)^6 (1 + 6/p + 1/p^2) \end{aligned}$$

and

$$\omega_\infty = 3 \iiint_{\{(t,u,v)\in\mathbb{R}^3 | 0\leqslant v\leqslant 1, |tv^2|, |v^2u|, |v(tv+u^2)|, |t(tv+u^2)|\leqslant 1\}} 1 \, \mathrm{d}t \, \mathrm{d}u \, \mathrm{d}v$$

Proof. Since S is split over $\mathbb{Q}$, we have $\mathrm{rk}\,(\mathrm{Pic}(\widetilde{S})) = 6$, and the expected exponent of $\log B$ is 5. Further, $\beta(S) = 1$. The computation of $c_{S,H}$ is done on the desingularization $\widetilde{S}$. For the computation of $\alpha(S)$, observe that the effective cone of $\widetilde{S}$ in $\mathrm{Pic}(\widetilde{S})$ is simplicial (it is generated by the exceptional curves $E_1, \dots, E_6$, and their number equals the rank of $\mathrm{Pic}(\widetilde{S})$), and

$$-K_{\widetilde{S}} = 4E_1 + 2E_2 + 3E_3 + 3E_4 + 2E_5 + 2E_6.$$

The calculation is analog to (Derenthal, 2005, Lemma 2) (see (Derenthal, 2006a) for its calculation in general). The constant $\omega_H(S)$ is computed as in (de la Bretèche and Browning, 2006, Lemma 1) and (Derenthal, 2005, Lemma 2).

4. The Universal Torsor

As explained above, the problem of counting rational points of bounded height on the surface S translates into a counting problem for certain integral points on the universal torsor, subject to coprimality and height inequalities. In the first part of this section, we describe these conditions in detail. They are obtained by a process of introducing new variables which are the greatest common divisors of other variables. Geometrically, this corresponds to the realization of $\widetilde{S}$ as a blow-up of $\mathbb{P}^2$ in five points.

In the second part, we prove our claims.

The universal torsor $\mathcal{T}_S$ of S is an open subset of the hypersurface in $\mathbb{A}^9 = \mathrm{Spec}\mathbb{Z}[\eta_1, \dots, \eta_6, \alpha_1, \alpha_2, \alpha_3]$ defined by the equation

$$T(\boldsymbol{\eta}, \boldsymbol{\alpha}) = \alpha_1^2\eta_2 + \alpha_2\eta_3\eta_5^2 + \alpha_3\eta_4\eta_6^2 = 0. \tag{5}$$

The projection $\Psi : \mathcal{T}_S \to S$ is defined by

$$(\Psi^*(x_i)) = (\eta^{(2,1,2,1,2,0)}\alpha_2, \eta^{(4,2,3,3,2,2)}, \eta^{(3,2,2,2,1,1)}\alpha_1, \eta^{(2,1,1,2,0,2)}\alpha_3, \alpha_2\alpha_3), \tag{6}$$

where we use the notation $\eta^{(n_1,n_2,n_3,n_4,n_5,n_6)} = \eta_1^{n_1}\eta_2^{n_2}\eta_3^{n_3}\eta_4^{n_4}\eta_5^{n_5}\eta_6^{n_6}$.

The coprimality conditions can be derived from the extended Dynkin diagram (see Figure 3). Two variables are allowed to have a common factor if and only if the corresponding divisors (E_i for η_i and A_i for α_i) intersect (i.e., are connected by an edge in the diagram). Furthermore, $\gcd(\alpha_1, \alpha_2, \alpha_3) > 1$ is allowed (corresponding to the fact that A_1, A_2, A_3 intersect in one point).

We will show below that there is a bijection between rational points on $S^0 \subset S$ and integral points on an open subset of $\mathcal{T}_S$, subject to these coprimality conditions.

We will later refer to

$$\text{coprimality between } \eta_i \text{ as in Figure 3,} \tag{7}$$
$$\gcd(\alpha_1, \eta_1\eta_3\eta_4\eta_5\eta_6) = 1, \tag{8}$$
$$\gcd(\alpha_2, \eta_1\eta_2\eta_3\eta_4\eta_6) = 1, \tag{9}$$
$$\gcd(\alpha_3, \eta_1\eta_2\eta_3\eta_4\eta_5) = 1. \tag{10}$$

To count the number of $\mathbf{x} \in S(\mathbb{Q})$ such that $H(\mathbf{x}) \leqslant B$, we must lift this condition to the universal torsor, i.e., $H(\Psi(\boldsymbol{\eta}, \boldsymbol{\alpha})) \leqslant B$. This is the same as

$$|\eta^{(2,1,2,1,2,0)}\alpha_2| \leqslant B, \qquad \dots, \qquad |\alpha_2\alpha_3| \leqslant B,$$

using the five monomials occurring in (6). These have no common factors, provided the coprimality conditions are fulfilled (direct verification).

It will be useful to write the height conditions as follows: let

$$X_0 = \left(\frac{\eta^{(4,2,3,3,2,2)}}{B}\right)^{1/3}, \quad X_1 = (B\eta^{(-1,-2,0,0,1,1)})^{1/3}, \quad X_2 = (B\eta^{(2,1,0,3,-2,4)})^{1/3}.$$

Then

$$|X_0^3| \leqslant 1 \tag{11}$$
$$|X_0^2(\alpha_1/X_1)| \leqslant 1 \tag{12}$$
$$|X_0^2(\alpha_2/X_2)| \leqslant 1, \quad |X_0(X_0(\alpha_2/X_2) + (\alpha_1/X_1)^2)| \leqslant 1,$$
$$|(\alpha_2/X_2)(X_0(\alpha_2/X_2) + (\alpha_1/X_1)^2)| \leqslant 1 \tag{13}$$

are equivalent to the five height conditions. Here we have used the torsor equation to eliminate α_3 because in our counting argument we will also use that α_3 is determined by the other variables.

We now prove the above claims.

LEMMA 4.1. *The map Ψ gives a bijection between the set of points $\mathbf{x}$ of $S^0(\mathbb{Q})$ such that $H(\mathbf{x}) \leqslant B$ and the set*

$$\mathcal{T}_1 := \left\{ (\boldsymbol{\eta}, \boldsymbol{\alpha}) \in \mathbb{Z}_{>0}^6 \times \mathbb{Z}^3 \;\middle|\; \begin{array}{l} \textit{equation } (5), \\ \textit{coprimality } (7), (8), (9), (10), \\ \textit{inequalities } (11),\ (12),\ (10) \textit{ hold} \end{array} \right\}$$

Proof. The map ψ of Lemma 2.1(3) induces a bijection

$$\psi_0 \colon (\eta_3, \alpha_1, \alpha_2) \mapsto (\eta_3^2\alpha_2, \eta_3^3, \eta_3^2\alpha_1, \eta_3\alpha_3, \alpha_2\alpha_3),$$

where $\alpha_3 := -(\eta_3\alpha_2 + \alpha_1^2)$, i.e.,

$$T_0 := \alpha_1^2 + \eta_3\alpha_2 + \alpha_3 = 0,$$

between

$$\{(\eta_3, \alpha_1, \alpha_2) \in \mathbb{Z}_{>0} \times \mathbb{Z}^2 \mid \gcd(\eta_3, \alpha_1, \alpha_2) = 1\} \quad \text{and} \quad S^0(\mathbb{Q}) \subset S(\mathbb{Q}).$$

The height function on $S^0(\mathbb{Q})$ is given by

$$H(\psi_0(\eta_3, \alpha_1, \alpha_2)) = \frac{\max(|\eta_3^2\alpha_2|, |\eta_3^3|, |\eta_3^2\alpha_1|, |\eta_3\alpha_3|, |\alpha_2\alpha_3|)}{\gcd(\eta_3^2\alpha_2, \eta_3^3, \eta_3^2\alpha_1, \eta_3\alpha_3, \alpha_2\alpha_3)}.$$

The derivation of the torsor equation from the map ψ_0 together with the coprimality conditions and the lifted height function is parallel to the blow-up process described in Lemma 2.2. More precisely, each line E_3, A_1, A_2 in $\mathbb{P}^2$ corresponds to a coordinate function η_3, α_1, α_2 vanishing in one of the lines; the blow-up of the intersection of two divisors gives an exceptional curve E_i, corresponding to the introduction of a new variable η_i as the greatest common divisor of two old variables. Two divisors are disjoint if and only if the corresponding variables are coprime. This is summarized in Table II.

TABLE II. Dictionary between gcd-process and blow-ups.

Variables, Equations	Geometry
variables	divisors
initial variables	coordinate lines
$\eta_3, \alpha_1, \alpha_2$	E_3, A_1, A_2
taking gcd of two variables	blowing up intersection of divisors
new gcd-variable	exceptional curve
$\eta_2, \eta_1, \eta_4, \eta_6, \eta_5$	E_2, E_1, E_4, E_6, E_5
extra variable	extra curve
α_3	A_3
starting relation	starting description
$\alpha_3 = -(\eta_3\alpha_2 + \alpha_1^2)$	$A_3 = \{\eta_3\alpha_2 + \alpha_1^2 = 0\}$
final relation	torsor equation
$\alpha_3\eta_4\eta_6^2 = -(\alpha_2\eta_3\eta_5^2 + \alpha_1^2\eta_2)$	$\alpha_1^2\eta_2 + \alpha_2\eta_3\eta_5^2 + \alpha_3\eta_4\eta_6^2 = 0$

This plan will now be implemented in five steps; at each step, the map

$$\psi_i \colon \mathbb{Z}_{>0}^{i+1} \times \mathbb{Z}^3 \to S^0(\mathbb{Q})$$

gives a bijection between:

- the set of all $(\eta_j, \alpha_1, \alpha_2, \alpha_3) \in \mathbb{Z}_{>0}^{i+} \times \mathbb{Z}^3$ satisfying certain coprimality conditions (described by the extended Dynkin diagram corresponding to the i-th blow-up of Lemma 2.2), an equation T_i,

$$H(\psi_i(\eta_j, \alpha_j)) = \frac{\max_k(|\psi_i(\eta_j, \alpha_j)_k|)}{\gcd(\psi_i(\eta_j, \alpha_j)_k)} \leqslant B,$$

- the set of all $\mathbf{x} \in S^0(\mathbb{Q})$ with $H(\mathbf{x}) \leqslant B$.

The steps are as follows:

1. Let $\eta_2 := \gcd(\eta_3, \alpha_1) \in \mathbb{Z}_{>0}$. Then

$$\eta_3 = \eta_2\eta_3', \quad \alpha_1 = \eta_2\alpha_1', \qquad \text{with } \gcd(\eta_3', \alpha_1') = 1.$$

Since $\eta_2 \mid \alpha_3$, we can write $\alpha_3 = \eta_2\alpha_3'$. Then $\alpha_3' = -(\eta_3'\alpha_2 + \eta_2\alpha_1'^2)$. After renaming the variables, we have

$$T_1 = \eta_2\alpha_1^2 + \eta_3\alpha_2 + \alpha_3 = 0$$

and

$$\psi_1\colon (\eta_2, \eta_3, \alpha_1, \alpha_2, \alpha_3) \mapsto (\eta_2\eta_3^2\alpha_2 : \eta_2^2\eta_3^3 : \eta_2^2\eta_3^2\alpha_1 : \eta_2\eta_3\alpha_3 : \alpha_2\alpha_3).$$

Here, we have eliminated the common factor η_2 which occurred in all five components of the image. Below, we repeat the corresponding transformation at each step.

2. Let $\eta_1 := \gcd(\eta_2, \eta_3) \in \mathbb{Z}_{>0}$. Then

$$\eta_2 = \eta_1\eta_2', \quad \eta_3 = \eta_1\eta_3', \qquad \text{with } \gcd(\eta_2', \eta_3') = 1.$$

As $\eta_1 \mid \alpha_3$, we write $\alpha_3 = \eta_1\alpha_3'$, and we obtain:

$$T_2 = \eta_2\alpha_1^2 + \eta_3\alpha_2 + \alpha_3 = 0$$

and

$$\psi_2\colon (\eta_1, \eta_2, \eta_3, \alpha_1, \alpha_2, \alpha_3) \mapsto \\ (\eta_1^2\eta_2\eta_3^2\alpha_2 : \eta_1^4\eta_2^2\eta_3^3 : \eta_1^3\eta_2^2\eta_3^2\alpha_1 : \eta_1^2\eta_2\eta_3\alpha_3 : \alpha_2\alpha_3).$$

3. Let $\eta_4 := \gcd(\eta_1, \alpha_3) \in \mathbb{Z}_{>0}$. Then

$$\eta_1 = \eta_4\eta_1', \quad \alpha_3 = \eta_4\alpha_3', \qquad \text{with } \gcd(\eta_1', \alpha_3') = 1.$$

We get after removing $'$ again:

$$T_3 = \eta_2\alpha_1^2 + \eta_3\alpha_2 + \eta_4\alpha_3 = 0$$

and

$$\psi_3\colon (\eta_1, \eta_2, \eta_3, \eta_4, \alpha_1, \alpha_2, \alpha_3) \mapsto \\ (\eta_1^2\eta_2\eta_3^2\eta_4\alpha_2 : \eta_1^4\eta_2^2\eta_3^3\eta_4^3 : \eta_1^3\eta_2^2\eta_3^2\eta_4^2\alpha_1 : \eta_1^2\eta_2\eta_3\eta_4^2\alpha_3 : \alpha_2\alpha_3).$$

4. Let $\eta_6 := \gcd(\eta_4, \alpha_3) \in \mathbb{Z}_{>0}$. Then
$$\eta_4 = \eta_6\eta_4', \qquad \alpha_3 = \eta_6\alpha_3', \qquad \text{with } \gcd(\eta_4', \alpha_3') = 1.$$
We obtain
$$T_4 = \eta_2\alpha_1^2 + \eta_3\alpha_2 + \eta_4\eta_6^2\alpha_3 = 0$$
and
$$\begin{aligned}&\psi_4\colon (\eta_1, \eta_2, \eta_3, \eta_4, \eta_6, \alpha_1, \alpha_2, \alpha_3) \mapsto \\ &\quad (\eta_1^2\eta_2\eta_3^2\eta_4\alpha_2 : \eta_1^4\eta_2^2\eta_3^3\eta_4^3\eta_6^2 : \eta_1^3\eta_2^2\eta_3^2\eta_4^2\eta_6\alpha_1 : \eta_1^2\eta_2\eta_3\eta_4^2\eta_6^2\alpha_3 : \alpha_2\alpha_3).\end{aligned}$$

5. The final step is $\eta_5 := \gcd(\eta_3, \alpha_2) \in \mathbb{Z}_{>0}$, we could have done it earlier (just as the blow-up of the intersection of E_3, A_2 in Lemma (6)). Then
$$\eta_3 = \eta_5\eta_3', \quad \alpha_2 = \eta_5\alpha_2', \qquad \text{with } \gcd(\eta_3', \alpha_2') = 1.$$
We get
$$T_5 = \eta_2\alpha_1^2 + \eta_3\eta_5\alpha_2 + \eta_4\eta_6^2\alpha_3 = 0$$
and
$$\begin{aligned}&\psi_5\colon (\eta_1, \eta_2, \eta_3, \eta_4, \eta_5, \eta_6, \alpha_1, \alpha_2, \alpha_3) \mapsto \\ &(\eta_1^2\eta_2\eta_3^2\eta_4\eta_5^2\alpha_2 : \eta_1^4\eta_2^2\eta_3^3\eta_4^3\eta_5^2\eta_6^2 : \eta_1^3\eta_2^2\eta_3^2\eta_4^2\eta_5\eta_6\alpha_1 : \eta_1^2\eta_2\eta_3\eta_4^2\eta_6^2\alpha_3 : \alpha_2\alpha_3)\end{aligned}$$

We observe that at each stage the coprimality conditions correspond to intersection properties of the respective divisors. The final result is summarized in Figure 3, which encodes data from (7), (8), (9), (10).

Note that ψ_5 is Ψ from (6). As mentioned above, $\gcd(\psi_5(\eta_j, \alpha_j)_k)$ (over all five components of the image) is trivial by the coprimality conditions of Figure 3. Therefore, $H(\psi_5(\boldsymbol{\eta}, \boldsymbol{\alpha})) \leqslant B$ is equivalent to (11), (12), (13).

Finally, T_5 is the torsor equation T (5).

5. Summations

In the first step, we estimate the number of $(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{Z}^3$ which fulfill the torsor equation T (5) and the height and coprimality conditions. For fixed (α_1, α_2), the torsor equation T has a solution α_3 if and only if the congruence
$$\alpha_1^2\eta_2 + \alpha_2\eta_3\eta_5^2 \equiv 0 \pmod{\eta_4\eta_6^2}$$
holds and the conditions on the height and coprimalities are fulfilled.

We have already written the height conditions so that they do not depend on α_3. For the coprimality, we must ensure that (9) and (10) are fulfilled.

As $\gcd(\eta_3\eta_5^2, \eta_4\eta_6^2) = 1$, we can find the multiplicative inverse c_1 of $\eta_3\eta_5^2$ modulo $\eta_4\eta_6^2$, so that

$$c_1\eta_3\eta_5^2 = 1 + c_2\eta_4\eta_6^2 \tag{14}$$

for a suitable c_2. Choosing

$$\alpha_2 = c_3\eta_4\eta_6^2 - c_1\alpha_1^2\eta_2, \tag{15}$$

$$\alpha_3 = c_2\alpha_1^2\eta_2 - c_3\eta_3\eta_5^2 \tag{16}$$

gives a solution of (5) for any $c_3 \in \mathbb{Z}$.

Without the coprimality conditions, the number of pairs (α_2, α_3) satisfying T and (13) would differ at most by $O(1)$ from $1/\eta_4\eta_6^2$ of the length of the interval described by (13). However, the coprimality conditions (9) and (10) impose further restrictions on the choice of c_3. A slight complication arises from the fact that because of T, some of the conditions are fulfilled automatically once $\boldsymbol{\eta}, \alpha_1$ satisfy (7) and (8).

Conditions (7) imply that the possibilities for a prime p to divide more than one of the η_i are very limited. We distinguish twelve cases, listed in Column 2 of Table III.

TABLE III. Coprimality conditions.

case	$p \mid \dots$	$p \mid \alpha_1$	$p \mid \alpha_2$	$p \mid \alpha_3$
0	–	allowed	allowed	allowed
i	η_1	restriction	restriction	restriction
ii	η_2	allowed	restriction	automatically
iii	η_3	restriction	restriction	automatically
iv	η_4	restriction	automatically	restriction
v	η_5	restriction	allowed	automatically
vi	η_6	restriction	automatically	allowed
vii	η_1, η_2	restriction	restriction	automatically
viii	η_1, η_3	restriction	restriction	automatically
ix	η_1, η_4	restriction	automatically	restriction
x	η_3, η_5	restriction	restriction	automatically
xi	η_4, η_6	restriction	automatically	restriction

In Columns 4 and 5, we have denoted the relevant information for the divisibility of α_2, α_3 by primes p which are divisors of the η_i in Column 2, but of no other η_j:

- “allowed” means that α_i may be divisible by p.
- “automatically” means that the conditions on the η_i and the other α_j imply that $p \nmid \alpha_i$. These two cases do not impose conditions on c_3 modulo p.

- "restriction" means that c_3 is not allowed to be in a certain congruence class modulo p in order to fulfill the condition that p must not divide α_i.

The information in the table is derived as follows:

- If $p \mid \eta_3$, then $p \nmid c_2$ from (14), and $p \nmid \alpha_1\eta_2$ because of (7), (8), so by (16), $p \nmid \alpha_3$ independently of the choice of c_3. Since $p \nmid \eta_4\eta_6^2$, we see from (15) that $p \mid \alpha_2$ for one in p subsequent choices of c_3 which we must therefore exclude. This explains cases *iii* and *viii*.
- In case *vii*, the same is true for α_2. More precisely, we see that we must exclude $c_3 \equiv 0 \pmod p$. By (16), $p \nmid c_3$ implies that $p \nmid \alpha_3$, so we do not need another condition on c_3.
- In case *i*, we see that $p \mid \alpha_2$ for one in p subsequent choices of c_3, and the same holds for α_3. However, in this case, p cannot divide α_2, α_3 for the same choice of c_3, as we can see by considering T: since $p \nmid \alpha_1^2\eta_2$, it is impossible that $p \mid \alpha_2, \alpha_3$. Therefore, we must exclude two out of p subsequent choices of p in order to fulfill $p \nmid \alpha_2, \alpha_3$.
- In the other cases, the arguments are similar.

The number of $(\alpha_2, \alpha_3) \in \mathbb{Z}^2$ subject to T, (9), (10), (13) equals the number of c_3 such that α_2, α_3 as in (15), (16) satisfy these conditions. This can be estimated as $1/\eta_4\eta_6^2$ of the interval described by (13), multiplied by a product of local factors whose values can be read off from Columns 2,4, 5 of Table III: the divisibility properties of η_i by p determine whether zero, one or two out of p subsequent values of c_3 have to be excluded. Different primes can be considered separately, and we define

$$\vartheta_{1,p} := \begin{cases} 1 - 2/p, & \text{case } i, \\ 1 - 1/p, & \text{cases } ii\text{–}iv,\ vi\text{–}xi, \\ 1, & \text{case } 0,\ v. \end{cases}$$

Let

$$\vartheta_1(\boldsymbol{\eta}) = \prod_p \vartheta_{1,p}$$

be the product of these local factors, and

$$g_1(u, v) = \int_{\{t \in \mathbb{R} \mid |tv^2|, |t(tv+u^2)|, |v(tv+u^2)| \leqslant 1\}} 1 \, \mathrm{d}t. \tag{17}$$

Let $\omega(n)$ denote the number of primes dividing n.

LEMMA 5.1. *For fixed* $(\boldsymbol{\eta}, \alpha_1) \in \mathbb{Z}_{>0}^6 \times \mathbb{Z}$ *as in* (7), (8), (11), (12), *the number of* $(\alpha_2, \alpha_3) \in \mathbb{Z}^2$ *satisfying* T, (9), (10), (13) *is*

$$\mathcal{N}_1(\boldsymbol{\eta}, \alpha_1) = \frac{\vartheta_1(\boldsymbol{\eta})X_2}{\eta_4\eta_6^2} g_1(\alpha_1/X_1, X_0) + O(2^{\omega(\eta_1\eta_2\eta_3\eta_4\eta_6)}).$$

The sum of error terms for all possible values of $(\boldsymbol{\eta}, \alpha_1)$ *is* $\ll B(\log B)^3$.

Proof. The number of c_3 such that the resulting α_2, α_3 satisfy (13) differs from $(X_2/\eta_4\eta_6^2)g_1(\alpha_1/X_1, X_0)$ by at most $O(1)$.

Each $\vartheta_{1,p} \neq 1$ corresponds to a congruence condition on c_3 imposed by one of the cases *i–iv*, *vi–xi*. For each congruence condition, the actual ratio of allowed c_3 can differ at most by $O(1)$ from the $\vartheta_{1,p}$. The total number of these primes p is

$$\omega(\eta_1\eta_2\eta_3\eta_4\eta_6) \ll 2^{\omega(\eta_1\eta_2\eta_3\eta_4\eta_6)},$$

which is independent of η_5 since any prime dividing only η_5 contributes a trivial factor (see case *v*).

Using the estimate (12) for α_1 in the first step and ignoring (7) (8), which can only increase the error term, we obtain:

$$\sum_{\boldsymbol{\eta}} \sum_{\alpha_1} 2^{\omega(\eta_1\eta_2\eta_3\eta_4\eta_6)} \leqslant \sum_{\boldsymbol{\eta}} \frac{B \cdot 2^{\omega(\eta_1\eta_2\eta_3\eta_4\eta_6)}}{\eta^{(3,2,2,2,1,1)}} \ll B(\log B)^3.$$

Here, we use $2^{\omega(n)} \ll_\epsilon n^\epsilon$ for the summations over $\eta_1, \eta_2, \eta_3, \eta_4$. For η_6, we employ

$$\sum_{n \leqslant x} 2^{\omega(n)} \ll x(\log x)$$

together with partial summation, contributing a factor $(\log B)^2$, while the summation over η_5 gives another factor $\log B$.

Next, we sum over all α_1 subject to the coprimality condition (8) and the height condition (12). Let

$$g_2(v) = \int_{\{u \in \mathbb{R} \mid |v^2 u| \leqslant 1\}} g_1(u, v)\, \mathrm{d}u \tag{18}$$

Similar to our discussion for α_2, α_3, the number of possible values for α_1 as in (12), while ignoring (8) for the moment, is $X_1 g_2(X_0) + O(1)$.

None of the coprimality conditions are fulfilled automatically, and only common factors with η_2 are allowed (see Column 3 of Table III). Therefore, each prime factor of $\eta_1\eta_3\eta_4\eta_5\eta_6$ reduces the number of allowed α_1 by a factor of $\vartheta_{2,p} = 1 - 1/p$ with an error of at most $O(1)$. For all other primes p, let $\vartheta_{2,p} = 1$, and let

$$\vartheta_2(\boldsymbol{\eta}) = \prod_p \vartheta_{2,p} \quad \text{and} \quad \vartheta(\boldsymbol{\eta}) = \begin{cases} \vartheta_1(\boldsymbol{\eta}) \cdot \vartheta_2(\boldsymbol{\eta}), & \text{(7) holds} \\ 0, & \text{otherwise.} \end{cases}$$

LEMMA 5.2. *For fixed* $\eta \in \mathbb{Z}_{>0}^6$ *as in* (7), (11), *the sum of* $\mathcal{N}_1(\boldsymbol{\eta}, \alpha_1)$ *over all* $\alpha_1 \in \mathbb{Z}$ *satisfying* (8), (12) *is*

$$\mathcal{N}_2(\boldsymbol{\eta}) := \frac{\vartheta(\boldsymbol{\eta})X_1X_2}{\eta_4\eta_6^2} g_2(X_0) + \mathcal{R}_2(\boldsymbol{\eta}),$$

where the sum of error terms $\mathcal{R}_2(\boldsymbol{\eta})$ *over all possible* $\boldsymbol{\eta}$ *is* $\ll B \log B$.

Proof. Let

$$\mathcal{N}(b_1, b_2) = \vartheta_1(\boldsymbol{\eta}) \cdot \#\{\alpha_1 \in [b_1, b_2] \mid \gcd(\alpha_1, \eta_1\eta_3\eta_4\eta_5\eta_6) = 1\}.$$

Using Möbius inversion, this is estimated as

$$\mathcal{N}(b_1, b_2) = \vartheta_1(\boldsymbol{\eta}) \cdot \vartheta_2(\boldsymbol{\eta}) \cdot (b_2 - b_1) + \mathcal{R}(b_1, b_2)$$

with $\mathcal{R}(b_1, b_2) = O(2^{\omega(\eta_1\eta_3\eta_4\eta_5\eta_6)})$. By partial summation,

$$\mathcal{N}_2(\boldsymbol{\eta}) = \frac{\vartheta(\boldsymbol{\eta})X_1X_2}{\eta_4\eta_6^2} g_2(X_0) + \mathcal{R}_2(\boldsymbol{\eta})$$

with

$$\mathcal{R}_2(\boldsymbol{\eta}) = \frac{-X_2}{\eta_4\eta_6^2} \int_{\{u \mid |X_0^2 u| \leqslant 1\}} (D_1 g_1)(u, X_0)\mathcal{R}(-X_1/X_0^2, X_1 u) \, \mathrm{d}u$$

where $D_1 g_1$ is the partial derivative of g_1 with respect to the first variable. Using the above bound for $\mathcal{R}(b_1, b_2)$, we obtain:

$$\mathcal{R}_2(\boldsymbol{\eta}) \ll \frac{X_2}{\eta_4\eta_6^2} 2^{\omega(\eta_1\eta_3\eta_4\eta_5\eta_6)}.$$

Summing this over all $\boldsymbol{\eta}$ as in (11) while ignoring (7) which can only enlarge the sum, we obtain:

$$\sum_{\boldsymbol{\eta}} \mathcal{R}_2(\boldsymbol{\eta}) \ll \sum_{\boldsymbol{\eta}} \frac{X_2 \cdot 2^{\omega(\eta_1\eta_3\eta_4\eta_5\eta_6)}}{\eta_4\eta_6^2 X_0^2} = \sum_{\boldsymbol{\eta}} \frac{B \cdot 2^{\omega(\eta_1\eta_3\eta_4\eta_5\eta_6)}}{\eta^{(2,1,2,2,2,2)}} \ll B \log B$$

In the first step, we use $X_0 \leqslant 1$.

Let

$$\Delta(n) = B^{-2/3} \sum_{\eta_i, \eta^{(4,2,3,3,2,2)} = n} \frac{\vartheta(\boldsymbol{\eta})X_1X_2}{\eta_4\eta_6^2} = \sum_{\eta_i, \eta^{(4,2,3,3,2,2)} = n} \frac{\vartheta(\boldsymbol{\eta})(\eta^{(4,2,3,3,2,2)})^{1/3}}{\eta^{(1,1,1,1,1,1)}}.$$

In view of Lemma 4.1, the number of rational points of bounded height on S^0 can be estimated by summing the result of Lemma 5.2 over all suitable $\boldsymbol{\eta}$. The error term is the combination of the error terms in Lemmas 5.1 and 5.2.

LEMMA 5.3. *We have*

$$N(S^0, -K_S, B) = B^{2/3} \sum_{n \leqslant B} \Delta(n) g_2((n/B)^{1/3}) + O(B(\log B)^3).$$

6. Completion of the Proof

We need an estimate for

$$M(t) := \sum_{n \leqslant t} \Delta(n).$$

Consider the Dirichlet series $F(s) := \sum_{n=1}^{\infty} \Delta(n) n^{-s}$. Using

$$F(s+1/3) = \sum_{\boldsymbol{\eta}} \frac{\vartheta(\boldsymbol{\eta})}{\eta_1^{4s+1}\eta_2^{2s+1}\eta_3^{3s+1}\eta_4^{3s+1}\eta_5^{2s+1}\eta_6^{2s+1}},$$

we write $F(s+1/3) = \prod_p F_p(s+1/3)$ as its Euler product. To obtain $F_p(s+1/3)$ for a prime p, we need to restrict this sum to the terms in which all η_i are powers of p. Note that $\vartheta(\boldsymbol{\eta})$ is non-zero if and only if the divisibility of η_i by p falls into one of the twelve cases described in Table III. The value of $\vartheta(\boldsymbol{\eta})$ only depends on these cases.

Writing $F_p(s+1/3) = \sum_{i=0}^{11} F_{p,i}(s+1/3)$, we have for example:

$$\begin{aligned}
F_{p,0}(s+1/3) &= 1,\\
F_{p,1}(s+1/3) &= \sum_{j=1}^{\infty} \frac{(1-1/p)(1-2/p)}{p^{j(4s+1)}} = \frac{(1-1/p)(1-2/p)}{p^{4s+1}-1},\\
F_{p,7}(s+1/3) &= \sum_{j,k=1}^{\infty} \frac{(1-1/p)^2}{p^{j(4s+1)}p^{k(2s+1)}} = \frac{(1-1/p)^2}{(p^{4s+1}-1)(p^{2s+1}-1)}.
\end{aligned}$$

The other cases are similar, giving

$$\begin{aligned}
F_p(s+1/3) &= 1 + \frac{1-1/p}{p^{4s+1}-1}\left((1-2/p) + \frac{1-1/p}{p^{2s+1}-1} + 2\frac{1-1/p}{p^{3s+1}-1}\right)\\
&\quad + \frac{1-1/p}{p^{2s+1}-1} + 2\frac{(1-1/p)^2}{p^{3s+1}-1} + 2\frac{1-1/p}{p^{2s+1}-1} + 2\frac{(1-1/p)^2}{(p^{2s+1}-1)^2}.
\end{aligned}$$

Defining

$$E(s) := \zeta(4s+1)\zeta(3s+1)^2\zeta(2s+1)^3 \quad \text{and} \quad G(s) := F(s+1/3)/E(s),$$

we see as in (Derenthal, 2005) that the residue of $F(s)t^s/s$ at $s = 1/3$ is

$$\mathrm{Res}(t) = \frac{3G(0)t^{1/3}Q_1(\log t)}{5! \cdot 4 \cdot 2 \cdot 3 \cdot 3 \cdot 2 \cdot 2}$$

for a monic $Q_1 \in \mathbb{R}[x]$ of degree 5. By Lemma 3.1, $\alpha(S) = \frac{1}{5!\cdot 4\cdot 2\cdot 3\cdot 3\cdot 2\cdot 2}$. By a Tauberian argument as in (Derenthal, 2005, Lemma 13):

LEMMA 6.1. *$M(t) = \mathrm{Res}(t) + O(t^{1/3-\delta})$ for some $\delta > 0$.*

By partial summation,

$$\sum_{n \leqslant B} \Delta(n) g_2((n/B)^{1/3}) = \alpha(S) G(0) B^{1/3} Q(\log B) \cdot 3 \int_0^1 g_2(v)\,\mathrm{d}v + O(B^{\frac{1}{3}-\delta})$$

for a monic polynomial Q of degree 5. We identify $\omega_H(S)$ from

$$G(0) = \prod_p \left(1 - \frac{1}{p}\right)^6 \left(1 + \frac{6}{p} + \frac{1}{p^2}\right), \quad \text{and} \quad \omega_\infty = 3 \int_0^1 g_2(v)\,\mathrm{d}v.$$

Together with Lemma 5.3, this completes the proof of Theorem 1.2.

7. Equations of Universal Torsors

The simplest universal torsors are those which can be realized as Zariski open subsets of the affine space. This happens if and only if the Del Pezzo surface is toric.

EXAMPLE 7.1. There are 20 types of singular Del Pezzo surfaces of degree $d \geqslant 3$ whose universal torsor is an open subset of a hypersurface in $\mathbb{A}^{13-d}$. For one example of each type[6], the equation defining the universal torsor is listed in Table IV. More details can be found in (Derenthal, 2006d).

EXAMPLE 7.2 (Cubic surface with $\mathbf{A}_1 + \mathbf{A}_3$ singularities). This surface has 7 lines, 4 additional variables correspond to exceptional curves of the desingularization. Its 9-dimensional universal torsor is a Zariski open subset of a complete intersection in

$$\mathbb{A}^{11} = \operatorname{Spec} \mathbb{Z}[\eta_0, \dots, \eta_3, \mu_0, \dots, \mu_6]$$

given by

$$\eta_1\eta_2\mu_1\mu_2 + \mu_4\mu_6 + \mu_3\mu_5 = 0 \quad \text{and} \quad \eta_0\eta_1\mu_2^2 + \eta_3\mu_5\mu_6 + \mu_0\mu_1 = 0.$$

See (Derenthal, 2006c) for more details.

[6] For the cubic $\mathbf{D}_4$ case, the universal torsor of a different example is calculated in (Hassett and Tschinkel, 2004, Section 4).

TABLE IV. Torsor equations.

degree	singularities	# of lines	defining equation
6	$\mathbf{A}_1$	3	$\eta_2\alpha_1 + \eta_3\alpha_2 + \eta_4\alpha_3$
6	$\mathbf{A}_2$	2	$\eta_2\alpha_1^2 + \eta_3\alpha_2 + \eta_4\alpha_3$
5	$\mathbf{A}_1$	7	$\eta_2\eta_6 + \eta_3\eta_7 + \eta_4\eta_8$
5	$\mathbf{A}_2$	4	$\eta_2\eta_5^2\eta_6 + \eta_3\alpha_1 + \eta_4\alpha_2$
5	$\mathbf{A}_3$	2	$\eta_1\alpha_1^2 + \eta_3\eta_4^2\alpha_2 + \eta_5\alpha_3$
5	$\mathbf{A}_4$	1	$\eta_1^2\eta_2\alpha_1^3 + \eta_4\alpha_2^2 + \eta_5\alpha_3$
4	$3\mathbf{A}_1$	6	$\eta_4\eta_5 + \eta_1\eta_6\eta_7 + \eta_8\eta_9$
4	$\mathbf{A}_2 + \mathbf{A}_1$	6	$\eta_5\eta_7 + \eta_1\eta_3\eta_9^2 + \eta_6\eta_8$
4	$\mathbf{A}_3$	5	$\eta_5\alpha + \eta_1\eta_4^2\eta_7 + \eta_3\eta_6^2\eta_8$
4	$\mathbf{A}_3 + \mathbf{A}_1$	3	$\eta_6\alpha_2 + \eta_7\alpha_1 + \eta_1\eta_3\eta_4^2\eta_5^3$
4	$\mathbf{A}_4$	3	$\eta_5\alpha_1 + \eta_1\alpha_2^2 + \eta_3\eta_4^2\eta_6^3\eta_7$
4	$\mathbf{D}_4$	2	$\eta_3\eta_5^2\alpha_2 + \eta_4\eta_6^2\alpha_3 + \eta_2\alpha_1^2$
4	$\mathbf{D}_5$	1	$\eta_3\alpha_1^2 + \eta_2\eta_6^2\alpha_3 + \eta_4\eta_5^2\alpha_2^3$
3	$\mathbf{D}_4$	6	$\eta_2\eta_5^2\eta_8 + \eta_3\eta_6^2\eta_9 + \eta_4\eta_7^2\eta_{10}$
3	$\mathbf{A}_3 + 2\mathbf{A}_1$	5	$\eta_4\eta_6^2\eta_{10} + \eta_1\eta_2\eta_7^2 + \eta_8\eta_9$
3	$2\mathbf{A}_2 + \mathbf{A}_1$	5	$\eta_3\eta_5\eta_7^2 + \eta_1\eta_6\eta_8 + \eta_9\eta_{10}$
3	$\mathbf{A}_4 + \mathbf{A}_1$	4	$\eta_1\eta_5\eta_8^2 + \eta_3\eta_4^2\eta_6^3\eta_9 + \eta_7\alpha$
3	$\mathbf{D}_5$	3	$\eta_2\eta_6^2\alpha_2 + \eta_4\eta_5^2\eta_7^3\eta_8 + \eta_3\alpha_1^2$
3	$\mathbf{A}_5 + \mathbf{A}_1$	2	$\eta_1^3\eta_2^2\eta_3\eta_7^4\eta_8 + \eta_5\alpha_1^2 + \eta_6\alpha_2$
3	$\mathbf{E}_6$	1	$\eta_4^2\eta_5\eta_7^3\alpha_3 + \eta_2\alpha_2^2 + \eta_1^2\eta_3\alpha_1^3$

There are examples of universal torsors which are not complete intersections, but have still been successfully used in the context of Manin's conjecture:

EXAMPLE 7.3 (Cayley cubic). The Cayley cubic surface

$$x_0x_1x_2 + x_0x_1x_3 + x_0x_2x_3 + x_1x_2x_3 = 0$$

(Figure 1) is a split singular cubic surface with four singularities $q_1, \dots, q_4$ of type $\mathbf{A}_1$ and nine lines. It is the blow-up of $\mathbb{P}^2$ in the 6 intersection points of 4 lines in general position. The universal torsor is an open subvariety of the variety in

$$\mathbb{A}^{13} = \operatorname{Spec}\mathbb{Z}[v_{12}, v_{13}, v_{14}, y_1, y_2, y_3, y_4, z_{12}, z_{13}, z_{14}, z_{23}, z_{24}, z_{25}]$$

defined by six equations of the form

$$z_{ik}z_{il}y_j + z_{jk}z_{jl}y_i = z_{ij}v_{ij}$$

and three equations of the form

$$v_{ij}v_{ik} = z_{il}^2y_jy_k - z_{jk}^2y_iy_l,$$

where $\{i, j, k, l\} = \{1, 2, 3, 4\}$ and

$$z_{ij} = z_{ji}, \quad v_{ij} = v_{ji}, \quad \text{and} \quad v_{ij} = -v_{kl}.$$

See (Derenthal, 2006c) for a proof. The variables y_i correspond to the four exceptional curves E_i obtained by blowing up q_i, z_{ij} correspond to the six lines m_{ij} through two of the singularities, and v_{ij} correspond to the other three lines ℓ_{ij}. The first six equations can be interpreted in connection with the projection from m_{ij}, and the other three equations are connected to the projection from ℓ_{ij}.

Upper and lower bounds of the expected order of magnitude have been established in (Heath-Brown, 2003).

EXAMPLE 7.4 (Smooth degree 5 Del Pezzo surface). The blow-up of $\mathbb{P}^2$ in

$$(1:0:0), \quad (0:1:0), \quad (0:0:1), \quad (1:1:1)$$

is a split smooth Del Pezzo surface of degree 5. Its universal torsor is an open subset of the variety defined by the following five equations in ten variables:

$$\begin{aligned}
\lambda_{12}\eta_2 - \lambda_{13}\eta_3 + \lambda_{14}\eta_4 &= 0\\
\lambda_{12}\eta_1 - \lambda_{23}\eta_3 + \lambda_{24}\eta_4 &= 0\\
\lambda_{13}\eta_1 - \lambda_{23}\eta_2 + \lambda_{34}\eta_4 &= 0\\
\lambda_{14}\eta_1 - \lambda_{24}\eta_2 + \lambda_{34}\eta_3 &= 0\\
\lambda_{12}\lambda_{34} - \lambda_{13}\lambda_{24} + \lambda_{14}\lambda_{23} &= 0
\end{aligned}$$

The asymptotic formula (3) has been established in (de la Bretèche, 2002).

To illustrate some of the difficulties in proving Conjecture 1.1 for a smooth split cubic surface, we now write down equations for its universal torsor (up to radical).

EXAMPLE 7.5 (Smooth cubic surfaces). Let S be the blow-up of $\mathbb{P}^2$ in

$$(1:0:0), \quad (0:1:0), \quad (0:0:1), \quad (1:1:1), \quad (1:a:b), \quad (1:c:d),$$

in general position. Conjecturally, the universal torsor is an open subset of the intersection of 81 quadrics in 27-dimensional space $\operatorname{Spec} \mathbb{Z}[\eta_i, \mu_{i,j}, \lambda_i]$, where

- $\eta_1, \ldots, \eta_6$ correspond to the preimages of the points,
- $\mu_{i,j}$ $(i < j \in \{1, \ldots, 6\})$ correspond to the 15 lines $m_{i,j}$ through two of the points,
- $\lambda_1, \ldots, \lambda_6$ correspond to the conics Q_i through five of the six points,

and relations arise from conic bundle structures on S. Batyrev and Popov proved that the above variables are indeed generators and that the relations give the universal torsor, up to radical (Batyrev and Popov, 2004).

We now write down these equations explicitly (see (Derenthal, 2006b) for more details). The 81 defining quadrics occur in sets of three. These 27 triples correspond to projections from the 27 lines on S. We use

$$E := (b-1)(c-1)-(a-1)(d-1) \quad \text{and} \quad F := bc - ad$$

to simplify the equations.

$$\begin{aligned}
q_{Q_1,1} &= -\eta_2\mu_{1,2} - \eta_3\mu_{1,3} + \eta_4\mu_{1,4}\\
q_{Q_1,2} &= -a\eta_2\mu_{1,2} - b\eta_3\mu_{1,3} + \eta_5\mu_{1,5}\\
q_{Q_1,2} &= -c\eta_2\mu_{1,2} - d\eta_3\mu_{1,3} + \eta_6\mu_{1,6}
\end{aligned}$$

$$\begin{aligned}
q_{Q_2,1} &= \eta_1\mu_{1,2} - \eta_3\mu_{2,3} + \eta_4\mu_{2,4}\\
q_{Q_2,2} &= \eta_1\mu_{1,2} - b\eta_3\mu_{2,3} + \eta_5\mu_{2,5}\\
q_{Q_2,3} &= \eta_1\mu_{1,2} - d\eta_3\mu_{2,3} + \eta_6\mu_{2,6}
\end{aligned}$$

$$\begin{aligned}
q_{Q_3,1} &= \eta_1\mu_{1,3} + \eta_2\mu_{2,3} + \eta_4\mu_{3,4}\\
q_{Q_3,2} &= \eta_1\mu_{1,3} + a\eta_2\mu_{2,3} + \eta_5\mu_{3,5}\\
q_{Q_3,3} &= \eta_1\mu_{1,3} + c\eta_2\mu_{2,3} + \eta_6\mu_{3,6}
\end{aligned}$$

$$\begin{aligned}
q_{Q_4,1} &= \eta_1\mu_{1,4} + \eta_2\mu_{2,4} + \eta_3\mu_{3,4}\\
q_{Q_4,2} &= (1-b)\eta_1\mu_{1,4} + (a-b)\eta_2\mu_{2,4} + \eta_5\mu_{4,5}\\
q_{Q_4,3} &= (1-d)\eta_1\mu_{1,4} + (c-d)\eta_2\mu_{2,4} + \eta_6\mu_{4,6}
\end{aligned}$$

$$\begin{aligned}
q_{Q_5,1} &= 1/b\eta_1\mu_{1,5} + a/b\eta_2\mu_{2,5} + \eta_3\mu_{3,5}\\
q_{Q_5,2} &= (1-b)/b\eta_1\mu_{1,5} + (a-b)/b\eta_2\mu_{2,5} + \eta_4\mu_{4,5}\\
q_{Q_5,3} &= (b-d)/b\eta_1\mu_{1,5} + F/b\eta_2\mu_{2,5} + \eta_6\mu_{5,6}
\end{aligned}$$

$$\begin{aligned}
q_{Q_6,1} &= 1/d\eta_1\mu_{1,6} + c/d\eta_2\mu_{2,6} + \eta_3\mu_{3,6}\\
q_{Q_6,2} &= (1-d)/d\eta_1\mu_{1,6} + (c-d)/d\eta_2\mu_{2,6} + \eta_4\mu_{4,6}\\
q_{Q_6,3} &= (b-d)/d\eta_1\mu_{1,6} + F/d\eta_2\mu_{2,6} + \eta_5\mu_{5,6}
\end{aligned}$$

$$\begin{aligned}
q_{m_{1,2},1} &= \mu_{4,5}\mu_{3,6} - \mu_{3,5}\mu_{4,6} + \mu_{3,4}\mu_{5,6}\\
q_{m_{1,2},2} &= (b-d)\mu_{3,5}\mu_{4,6} + (d-1)\mu_{3,4}\mu_{5,6} + \eta_2\lambda_1\\
q_{m_{1,2},3} &= F\mu_{3,5}\mu_{4,6} + a(d-c)\mu_{3,4}\mu_{5,6} + \eta_1\lambda_2
\end{aligned}$$

$$\begin{aligned}
q_{m_{1,3},1} &= \mu_{4,5}\mu_{2,6} - \mu_{2,5}\mu_{4,6} + \mu_{2,4}\mu_{5,6}\\
q_{m_{1,3},2} &= (c-a)\mu_{2,5}\mu_{4,6} + (1-c)\mu_{2,4}\mu_{5,6} + \eta_3\lambda_1
\end{aligned}$$

$$q_{m_{1,3},3} = -F\mu_{2,5}\mu_{4,6} + b(c-d)\mu_{2,4}\mu_{5,6} + \eta_1\lambda_3$$

$$\begin{aligned}
q_{m_{2,3},1} &= \mu_{4,5}\mu_{1,6} - \mu_{1,5}\mu_{4,6} + \mu_{1,4}\mu_{5,6}\\
q_{m_{2,3},2} &= (a-c)\mu_{1,5}\mu_{4,6} + a(c-1)\mu_{1,4}\mu_{5,6} + \eta_3\lambda_2\\
q_{m_{2,3},3} &= (b-d)\mu_{1,5}\mu_{4,6} + b(d-1)\mu_{1,4}\mu_{5,6} + \eta_2\lambda_3
\end{aligned}$$

$$\begin{aligned}
q_{m_{1,4},1} &= \mu_{3,5}\mu_{2,6} - \mu_{2,5}\mu_{3,6} + \mu_{2,3}\mu_{5,6}\\
q_{m_{1,4},2} &= -E\mu_{2,5}\mu_{3,6} + (b-1)(c-1)\mu_{2,3}\mu_{5,6} + \eta_4\lambda_1\\
q_{m_{1,4},3} &= -F\mu_{2,5}\mu_{3,6} + bc\mu_{2,3}\mu_{5,6} + \eta_1\lambda_4
\end{aligned}$$

$$\begin{aligned}
q_{m_{2,4},1} &= \mu_{3,5}\mu_{1,6} - \mu_{1,5}\mu_{3,6} + \mu_{1,3}\mu_{5,6}\\
q_{m_{2,4},2} &= E\mu_{1,5}\mu_{3,6} + (a-b)(c-1)\mu_{1,3}\mu_{5,6} + \eta_4\lambda_2\\
q_{m_{2,4},3} &= (b-d)\mu_{1,5}\mu_{3,6} - b\mu_{1,3}\mu_{5,6} + \eta_2\lambda_4
\end{aligned}$$

$$\begin{aligned}
q_{m_{3,4},1} &= \mu_{2,5}\mu_{1,6} - \mu_{1,5}\mu_{2,6} + \mu_{1,2}\mu_{5,6}\\
q_{m_{3,4},2} &= -E\mu_{1,5}\mu_{2,6} + (a-b)(1-d)\mu_{1,2}\mu_{5,6} + \eta_4\lambda_3\\
q_{m_{3,4},3} &= (c-a)\mu_{1,5}\mu_{2,6} + a\mu_{1,2}\mu_{5,6} + \eta_3\lambda_4
\end{aligned}$$

$$\begin{aligned}
q_{m_{1,5},1} &= \mu_{3,4}\mu_{2,6} - \mu_{2,4}\mu_{3,6} + \mu_{2,3}\mu_{4,6}\\
q_{m_{1,5},2} &= -E\mu_{2,4}\mu_{3,6} + (a-c)(1-b)\mu_{2,3}\mu_{4,6} + \eta_5\lambda_1\\
q_{m_{1,5},3} &= (d-c)\mu_{2,4}\mu_{3,6} + c\mu_{2,3}\mu_{4,6} + \eta_1\lambda_5
\end{aligned}$$

$$\begin{aligned}
q_{m_{2,5},1} &= \mu_{3,4}\mu_{1,6} - \mu_{1,4}\mu_{3,6} + \mu_{1,3}\mu_{4,6}\\
q_{m_{2,5},2} &= aE\mu_{1,4}\mu_{3,6} + (a-b)(c-a)\mu_{1,3}\mu_{4,6} + \eta_5\lambda_2\\
q_{m_{2,5},3} &= (1-d)\mu_{1,4}\mu_{3,6} - \mu_{1,3}\mu_{4,6} + \eta_2\lambda_5
\end{aligned}$$

$$\begin{aligned}
q_{m_{3,5},1} &= \mu_{2,4}\mu_{1,6} - \mu_{1,4}\mu_{2,6} + \mu_{1,2}\mu_{4,6}\\
q_{m_{3,5},2} &= -bE\mu_{1,4}\mu_{2,6} + (a-b)(b-d)\mu_{1,2}\mu_{4,6} + \eta_5\lambda_3\\
q_{m_{3,5},3} &= (c-1)\mu_{1,4}\mu_{2,6} + \mu_{1,2}\mu_{4,6} + \eta_3\lambda_5
\end{aligned}$$

$$\begin{aligned}
q_{m_{4,5},1} &= \mu_{2,3}\mu_{1,6} - \mu_{1,3}\mu_{2,6} + \mu_{1,2}\mu_{3,6}\\
q_{m_{4,5},2} &= b(c-a)\mu_{1,3}\mu_{2,6} + a(b-d)\mu_{1,2}\mu_{3,6} + \eta_5\lambda_4\\
q_{m_{4,5},3} &= (c-1)\mu_{1,3}\mu_{2,6} + (1-d)\mu_{1,2}\mu_{3,6} + \eta_4\lambda_5
\end{aligned}$$

$$\begin{aligned}
q_{m_{1,6},1} &= \mu_{3,4}\mu_{2,5} - \mu_{2,4}\mu_{3,5} + \mu_{2,3}\mu_{4,5}\\
q_{m_{1,6},2} &= -E\mu_{2,4}\mu_{3,5} + (a-c)(1-d)\mu_{2,3}\mu_{4,5} + \eta_6\lambda_1\\
q_{m_{1,6},3} &= (b-a)\mu_{2,4}\mu_{3,5} + a\mu_{2,3}\mu_{4,5} + \eta_1\lambda_6
\end{aligned}$$

$$\begin{aligned}
q_{m_{2,6},1} &= \mu_{3,4}\mu_{1,5} - \mu_{1,4}\mu_{3,5} + \mu_{1,3}\mu_{4,5}\\
q_{m_{2,6},2} &= cE\mu_{1,4}\mu_{3,5} + (a-c)(d-c)\mu_{1,3}\mu_{4,5} + \eta_6\lambda_2
\end{aligned}$$

$$\begin{aligned}
q_{m_{2,6},3} &= (1-b)\mu_{1,4}\mu_{3,5} - \mu_{1,3}\mu_{4,5} + \eta_2\lambda_6
\end{aligned}$$

$$\begin{aligned}
q_{m_{3,6},1} &= \mu_{2,4}\mu_{1,5} - \mu_{1,4}\mu_{2,5} + \mu_{1,2}\mu_{4,5}\\
q_{m_{3,6},2} &= -dE\mu_{1,4}\mu_{2,5} + (d-b)(d-c)\mu_{1,2}\mu_{4,5} + \eta_6\lambda_3\\
q_{m_{3,6},3} &= (a-1)\mu_{1,4}\mu_{2,5} + \mu_{1,2}\mu_{4,5} + \eta_3\lambda_6
\end{aligned}$$

$$\begin{aligned}
q_{m_{4,6},1} &= \mu_{2,3}\mu_{1,5} - \mu_{1,3}\mu_{2,5} + \mu_{1,2}\mu_{3,5}\\
q_{m_{4,6},2} &= d(c-a)\mu_{1,3}\mu_{2,5} + c(b-d)\mu_{1,2}\mu_{3,5} + \eta_6\lambda_4\\
q_{m_{4,6},3} &= (a-1)\mu_{1,3}\mu_{2,5} + (1-b)\mu_{1,2}\mu_{3,5} + \eta_4\lambda_6
\end{aligned}$$

$$\begin{aligned}
q_{m_{5,6},1} &= \mu_{2,3}\mu_{1,4} - \mu_{1,3}\mu_{2,4} + \mu_{1,2}\mu_{3,4}\\
q_{m_{5,6},2} &= d(c-1)\mu_{1,3}\mu_{2,4} + c(1-d)\mu_{1,2}\mu_{3,4} + \eta_6\lambda_5\\
q_{m_{5,6},3} &= b(a-1)\mu_{1,3}\mu_{2,4} + a(1-b)\mu_{1,2}\mu_{3,4} + \eta_5\lambda_6
\end{aligned}$$

$$\begin{aligned}
q_{E_1,1} &= (d-b)/E\mu_{1,2}\lambda_2 + (c-a)/E\mu_{1,3}\lambda_3 + \mu_{1,4}\lambda_4\\
q_{E_1,2} &= (d-1)/E\mu_{1,2}\lambda_2 + (c-1)/E\mu_{1,3}\lambda_3 + \mu_{1,5}\lambda_5\\
q_{E_1,3} &= (b-1)/E\mu_{1,2}\lambda_2 + (a-1)/E\mu_{1,3}\lambda_3 + \mu_{1,6}\lambda_6
\end{aligned}$$

$$\begin{aligned}
q_{E_2,1} &= F/E\mu_{1,2}\lambda_1 + (c-a)/E\mu_{2,3}\lambda_3 + \mu_{2,4}\lambda_4\\
q_{E_2,2} &= (c-d)/E\mu_{1,2}\lambda_1 + (c-1)/E\mu_{2,3}\lambda_3 + \mu_{2,5}\lambda_5\\
q_{E_2,3} &= (a-b)/E\mu_{1,2}\lambda_1 + (a-1)/E\mu_{2,3}\lambda_3 + \mu_{2,6}\lambda_6
\end{aligned}$$

$$\begin{aligned}
q_{E_3,1} &= F/E\mu_{1,3}\lambda_1 + (b-d)/E\mu_{2,3}\lambda_2 + \mu_{3,4}\lambda_4\\
q_{E_3,2} &= (c-d)/E\mu_{1,3}\lambda_1 + (1-d)/E\mu_{2,3}\lambda_2 + \mu_{3,5}\lambda_5\\
q_{E_3,3} &= (a-b)/E\mu_{1,3}\lambda_1 + (1-b)/E\mu_{2,3}\lambda_2 + \mu_{3,6}\lambda_6
\end{aligned}$$

$$\begin{aligned}
q_{E_4,1} &= F/(a-c)\mu_{1,4}\lambda_1 + (b-d)/(a-c)\mu_{2,4}\lambda_2 + \mu_{3,4}\lambda_3\\
q_{E_4,2} &= c/(a-c)\mu_{1,4}\lambda_1 + 1/(a-c)\mu_{2,4}\lambda_2 + \mu_{4,5}\lambda_5\\
q_{E_4,3} &= a/(a-c)\mu_{1,4}\lambda_1 + 1/(a-c)\mu_{2,4}\lambda_2 + \mu_{4,6}\lambda_6
\end{aligned}$$

$$\begin{aligned}
q_{E_5,1} &= (d-c)/(c-1)\mu_{1,5}\lambda_1 + (d-1)/(c-1)\mu_{2,5}\lambda_2 + \mu_{3,5}\lambda_3\\
q_{E_5,2} &= -c/(c-1)\mu_{1,5}\lambda_1 - 1/(c-1)\mu_{2,5}\lambda_2 + \mu_{4,5}\lambda_4\\
q_{E_5,3} &= -1/(c-1)\mu_{1,5}\lambda_1 - 1/(c-1)\mu_{2,5}\lambda_2 + \mu_{5,6}\lambda_6
\end{aligned}$$

$$\begin{aligned}
q_{E_6,1} &= (b-a)/(a-1)\mu_{1,6}\lambda_1 + (b-1)/(a-1)\mu_{2,6}\lambda_2 + \mu_{3,6}\lambda_3\\
q_{E_6,2} &= -a/(a-1)\mu_{1,6}\lambda_1 - 1/(a-1)\mu_{2,6}\lambda_2 + \mu_{4,6}\lambda_4\\
q_{E_6,3} &= -1/(a-1)\mu_{1,6}\lambda_1 - 1/(a-1)\mu_{2,6}\lambda_2 + \mu_{5,6}\lambda_5
\end{aligned}$$

In general, the dimension k of the ambient space $\mathbb{A}^k$ of the universal torsor is at least as large as the number of lines on the surface plus the number of exceptional curves of its desingularization, while the dimension of the universal torsor only depends on the degree of the surface, so that the number of equations must grow with k.

Heuristically, the complexity of universal torsors should be dictated by the following considerations:

- The dimension of the universal torsor of split Del Pezzo surfaces S is $12 - d$, where d is the degree of S.
- For smooth Del Pezzo surfaces, the number of lines is bigger in smaller degrees (e.g., 10 lines in degree 5, and 27 lines in degree 3).
- Singular surfaces have less lines than smooth surfaces.
- The number of lines is higher in cases with "few mild" singularities (e.g., for cubics: $\mathbf{A}_1$ with 21 lines, $\mathbf{A}_2$ with 15 lines), while it is low for "bad" singularities (e.g., 1 for the $\mathbf{E}_6$ cubic, 2 for the $\mathbf{A}_5 + \mathbf{A}_1$ cubic).

Therefore, we expect universal torsors over surfaces which have low degree, are smooth or have mild singularities to be more complex than torsors over surfaces in large degree, or with complicated singularities.

Acknowledgements

Part of this work was done while the authors were visiting the CRM at the Université de Montréal during the special year on *Analysis in Number Theory*. We are grateful for the invitation and ideal working conditions.

References

Batyrev, V. V. and Manin, Y. I. (1990) Sur le nombre des points rationnels de hauteur borné des variétés algébriques, *Math. Ann.* **286**, 27–43.

Batyrev, V. V. and Popov, O. N. (2004) The Cox ring of a del Pezzo surface, In *Arithmetic of higher-dimensional algebraic varieties*, Vol. 226 of *Progr. Math.*, pp. 85–103, Boston, MA, Birkhäuser Boston.

Batyrev, V. V. and Tschinkel, Y. (1998) Manin's conjecture for toric varieties, *J. Algebraic Geom.* **7**, 15–53.

Birch, B. J. (1961/1962) Forms in many variables, *Proc. Roy. Soc. Ser. A* **265**, 245–263.

Browning, T. D. (2004) The density of rational points on a certain singular cubic surface, arXiv:math.NT/0404245.

Browning, T. D. (2005) An overview of Manin's conjecture for del Pezzo surfaces, arXiv:math.NT/0511041.

Chambert-Loir, A. and Tschinkel, Y. (2002) On the distribution of points of bounded height on equivariant compactifications of vector groups, *Invent. Math.* **148**, 421–452.

Colliot-Thélène, J.-L. and Sansuc, J.-J. (1987) La descente sur les variétés rationnelles. II, *Duke Math. J.* **54**, 375–492.

Coray, D. F. and Tsfasman, M. A. (1988) Arithmetic on singular Del Pezzo surfaces, *Proc. London Math. Soc. (3)* **57**, 25–87.

de la Bretèche, R. (1998) Sur le nombre de points de hauteur bornée d'une certaine surface cubique singulière, In *Nombre et répartition de points de hauteur bornée*, Vol. 251 of *Astérisque*, Paris, 1996, pp. 51–77, Paris, Société Mathématique de France.

de la Bretèche, R. (2002) Nombre de points de hauteur bornée sur les surfaces de del Pezzo de degré 5, *Duke Math. J.* **113**, 421–464.

de la Bretèche, R. and Browning, T. D. (2005) On Manin's conjecture for singular del Pezzo surfaces of degree four, II, arXiv:math.NT/0502510.

de la Bretèche, R. and Browning, T. D. (2006) On Manin's conjecture for singular del Pezzo surfaces of degree four, I, *Mich. Math. J.*, to appear.

de la Bretèche, R., Browning, T. D., and Derenthal, U. (2005) On Manin's conjecture for a certain singular cubic surface, *Ann. Sci. École Norm. Sup. (4)*, to appear.

de la Bretèche, R. and Fouvry, É. (2004) L'éclaté du plan projectif en quatre points dont deux conjugués, *J. Reine Angew. Math.* **576**, 63–122.

Derenthal, U. (2005) Manin's conjecture for a certain singular cubic surface, arXiv:math.NT/0504016.

Derenthal, U. (2006a) On a constant arising in Manin's conjecture for Del Pezzo surfaces.

Derenthal, U. (2006b) On the Cox ring of Del Pezzo surfaces, arXiv:math.AG/0603111.

Derenthal, U. (2006c) On the Cox ring of singular Del Pezzo surfaces.

Derenthal, U. (2006d) Singular Del Pezzo surfaces whose universal torsors are hypersurfaces, math.AG/0604194.

Dolgachev, I. (2003) *Lectures on invariant theory*, Vol. 296 of *London Mathematical Society Lecture Note Series*, Cambridge, Cambridge University Press.

Franke, J., Manin, Y. I., and Tschinkel, Y. (1989) Rational points of bounded height on Fano varieties, *Invent. Math.* **95**, 421–435.

Hartshorne, R. (1977) *Algebraic geometry*, New York, Springer-Verlag, Graduate Texts in Mathematics, No. 52.

Hassett, B. and Tschinkel, Y. (2004) Universal torsors and Cox rings, In *Arithmetic of higher-dimensional algebraic varieties*, Vol. 226 of *Progr. Math.*, Palo Alto, CA, 2002, pp. 149–173, Boston, MA, Birkhäuser Boston.

Heath-Brown, D. R. (2003) The density of rational points on Cayley's cubic surface, In *Proceedings of the Session in Analytic Number Theory and Diophantine Equations*, Vol. 360 of *Bonner Math. Schriften*, Bonn, p. 33, Univ. Bonn.

Heath-Brown, D. R. (2006) Article in this volume.

Hu, Y. and Keel, S. (2000) Mori dream spaces and GIT, *Michigan Math. J.* **48**, 331–348, Dedicated to William Fulton on the occasion of his 60th birthday.

Manin, Y. I. (1986) *Cubic forms. Algebra, geometry, arithmetic*, Vol. 4 of *North-Holland Mathematical Library*, Amsterdam, North-Holland Publishing Co., 2nd edition.

Peyre, E. (1995) Hauteurs et mesures de Tamagawa sur les variétés de Fano, *Duke Math. J.* **79**, 101–218.

Peyre, E. (1998) Terme principal de la fonction zêta des hauteurs et torseurs universels, In *Nombre et répartition de points de hauteur bornée*, Vol. 251 of *Astérisque*, Paris, 1996, pp. 259–298, Paris, Société Mathématique de France.

Salberger, P. (1998) Tamagawa measures on universal torsors and points of bounded height on Fano varieties, In *Nombre et répartition de points de hauteur bornée*, Vol. 251 of *Astérisque*, Paris, 1996, pp. 91–258, Paris, Société Mathématique de France.

AN INTRODUCTION TO THE LINNIK PROBLEMS

W. Duke
UCLA

Abstract. This paper is a slightly enlarged version of a series of lectures on the Linnik problems given at the SMS–NATO ASI 2005 Summer School on Equidistribution in Number Theory.

Key words: Linnik problem, half-integral weight, CM points

1. Introduction

In these lectures I will discuss the classical Linnik problems about the distribution of lattice points on a sphere and analogous hyperbolic problems associated to binary quadratic forms. These problems were introduced by Linnik and are discussed in his book *Ergodic Properties of Algebraic Fields*. In (Linnik, 1968), Linnik applied an intricate ergodic method to solve them subject to a certain condition. In (Iwaniec, 1987), Iwaniec made a breakthrough in the theory of modular forms of half-integral weight that allowed the Linnik problems to be solved unconditionally using more traditional modular forms methods (Duke, 1988). These methods have since been much further developed in the more general context of subconvexity estimates for L-functions, where they have far-ranging implications/applications. My main purpose is to give an exposition of the original modular forms approach emphasizing the original ideas, which have an intuitive appeal. I will only introduce briefly the connection with L-functions. Recently there has been striking progress by a number of mathematicians in the analytic theory of L-functions in connection with various equidistribution problems. Hopefully, these lectures will provide some background for these developments, and serve as a rough guide to help those interested in pursuing details. An excellent exposition of many of the topics treated here is (Sarnak, 1990).

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 197–216.

2. The Linnik Problems

2.1. THE SPHERE

Consider the lattice points $\alpha \in \mathbb{Z}^3$ with $|\alpha|^2 = x_1^2 + x_2^2 + x_3^2$, for $\alpha = (x_1, x_2, x_3)$. The set $\Omega_n = \{x = \alpha/|\alpha|; \alpha \in \mathbb{Z}^3; |\alpha|^2 = n\}$ for $n \in \mathbb{Z}^+$ lies on the unit sphere S^2. By a classical result of Legendre Ω_n is non-empty iff $n \neq 4^a(8b + 7)$ for a and b integers, a non-negative. Linnik asked whether the set Ω_n subject to the condition that $n \equiv 1, 2, 3, 5, 6 \pmod 8$ is uniformly distributed with respect to (normalized) Lebesgue measure $d\sigma$ on S^2; is it the case given a reasonable subset of S^2 that the proportion of points in it from Ω_n approaches the measure of the set as $n \to \infty$? Linnik was able to prove this using his “ergodic method” but subject to the condition required by the method that the Legendre symbol $(n/p) = 1$ for a fixed odd prime p. An advance made by Iwaniec in the estimation of Fourier coefficients of cusp forms of half-integral weight later allowed this condition to be removed. To state this, it is convenient to couch the uniform distribution property in terms of the approximation of the integral of a test function by “Riemann sums.” For simplicity I will restrict attention here to the most interesting case where n is square-free.

THEOREM A. *Suppose that $f \in C^\infty(S^2)$. Then, as $n \to \infty$ with n square-free and $n \not\equiv 7 \pmod 8$,*

$$\frac{1}{\#\Omega_n} \sum_{x \in \Omega_n} f(x) \to \int_{S^2} f \, d\sigma.$$

I will spend most of the lectures explaining, modulo many technical details, the proof of this result. It should be pointed out that one may ask the same question about the lattice points on a ellipsoid given by a positive definite integral ternary quadratic form. Then, most of the interest shifts to the question of characterizing by mean of congruences those integers n that are represented by the form. For square-free n, the analytic techniques used to prove Theorem A apply directly, but for general n the issue becomes quite delicate (see e.g. (Duke and Schulze-Pillot, 1990; Duke, 1997)).

2.2. CM POINTS

Another problem introduced by Linnik concerns the distribution of roots of integral quadratic equations with a large negative discriminant. Here the appropriate setting is $\pm\Gamma\backslash\mathcal{H}$, where $\mathcal{H}$ is the upper half-plane and $\Gamma = \mathrm{SL}(2, \mathbb{Z})$ is the modular group. The quadratic equations are best introduced via positive definite binary quadratic forms

$$Q = Q(x, y) = ax^2 + bxy + cy^2, \quad d = b^2 - 4ac = \operatorname{disc} Q < 0$$

with $a, b, c \in \mathbb{Z}$, $a > 0$. After Gauss, there are only finitely many Γ-equivalence classes of such forms with a given d (see (Cox, 1989)).

For a given $Q = ax^2+bxy+cy^2$ with disc $Q = d$, the root of $ax^2+bx+c = 0$

$$z_Q = \frac{-b + \sqrt{d}}{2a} \in \mathcal{H}$$

associated to Q is called a CM point. It is readily shown that the orbit γz_Q runs over the roots of the forms equivalent to Q, where $\gamma \in \Gamma$ acts as a linear fractional map. Let $\mathcal{F}$ denote the standard fundamental domain for Γ:

$$\mathcal{F} = \left\{z \in \mathcal{H}; -\tfrac{1}{2} \leq \operatorname{Re} z \leq 0 \text{ and } |z| \geq 1 \text{ or } 0 < \operatorname{Re} z < \tfrac{1}{2} \text{ and } |z| > 1\right\}.$$

We shall write $\Lambda_d = \{z_Q \in \mathcal{F}; \operatorname{disc} Q = d\}$. For every $d \equiv 0, 1 (\operatorname{mod} 4)$ we can find a (principal) Q with disc $Q = d$ and associated $z_Q \in \mathcal{F}$:

$$\begin{aligned} d \equiv 0(4) \ &: \ x^2 - \frac{d}{4}y^2, \quad z_d = \frac{\sqrt{d}}{2}, \\ d \equiv 1(4) \ &: \ x^2 + xy - \frac{d-1}{y}y^2, \quad z_d = \frac{-1 + \sqrt{d}}{2}. \end{aligned}$$

It is convenient to define by convention a sum over Λ_d to mean that a summand should be weighted by $\frac{1}{2}$ if $Q = a(x^2+y^2)$ and by $\frac{1}{3}$ if $Q = a(x^2+xy+y^2)$ to account for the automorphs in Γ. In particular, $H(d) = \sum_{\Lambda_d} 1$ is called the Hurwitz class number. Recall that d is said to be fundamental when it equals the discriminant of $\mathbb{Q}(\sqrt{d})$. In this case, when $d < -4$, $H(d)$ equals to the class number $h(d)$ of $\mathbb{Q}(\sqrt{d})$ Generally I will only be concerned with fundamental discriminants.

A $\mathrm{PSL}(2, \mathbb{R})$-invariant measure for $\mathcal{H}$ is given by $dx\, dy/y^2$ and

$$\iint_{\mathcal{F}} dx\, dy/y^2 = \pi/3.$$

Let us denote by $d\mu = (3/\pi)\, dx\, dy/y^2$ the normalized invariant measure. The second Linnik problem concerns the distribution of the $z_Q \in \mathcal{F}$ as $d \to -\infty$.

THEOREM B. *Suppose that $f \in C^\infty(\mathcal{H})$ is Γ-invariant and bounded on $\mathcal{H}$. Then, as $d \to -\infty$ with d a fundamental discriminant,*

$$\frac{1}{\#\Lambda_d} \sum_{z \in \Lambda_d} f(z) \to \int_{\Gamma \backslash \mathcal{H}} f\, d\mu.$$

The proof of this result is quite analogous to that of Theorem A but requires more machinery. The main reason for this is the fact that $\Gamma \backslash \mathcal{H}$ is non-compact.

There is a parallel result one can obtain for indefinite forms as $d \to +\infty$, namely the uniform distribution of closed geodesics on $\pm\Gamma\backslash\mathcal{H}$ when grouped by discriminant. In fact, the proof of Theorem B yields this result as well. This problem is in fact a revealing paradigm for more general situations in which infinite unit groups exist (see. e.g. (Cohen, 2005) and references given there).

3. Holomorphic Modular Forms of Half-Integral Weight

This subject is based on the properties of the Jacobi theta series

$$\theta(z) = \sum_{n\in\mathbb{Z}} e(n^2 z),$$

which has a product representation via the Jacobi triple product formula: write $q = e(z)$

$$\theta(z) = \prod_{n=1}^{\infty}(1-q^{2n})(1+q^{2n-1})^2.$$

This remarkable function satisfies for $\gamma\in\Gamma_0(4)$, where $\Gamma_0(N) = \{\gamma \in \mathrm{SL}(2,\mathbb{Z})$: $c \equiv 0(N)\}$, the transformation formula

$$\theta(\gamma z) = j(\gamma, z)\theta(z),$$

where $j(\gamma, z) = (c/d)\varepsilon_d^{-1}(cz+d)^{1/2}$, with (c/d) the (extended) Legendre symbol, $\varepsilon_d = \begin{cases} 1, & d \equiv 1(4) \\ i, & d \equiv 3(4) \end{cases}$ and $z^{1/2} = |z|^{1/2}\exp(\frac{1}{2}i\arg z)$, with $-\pi < \arg z \leq \pi$, (see (Shimura, 1973)). Actually, Jacobi studied $\theta(z/2)$, whose relevant group is conjugate to $\Gamma_0(4)$, namely $\Gamma(2)$.

For $k \in \frac{1}{2}\mathbb{Z}^+$ and $N \equiv 0(4)$ if $2k$ is odd, a holomorphic modular form of weight k for $\Gamma_0(N)$ is a holomorphic function on $\mathcal{H}$ sit for $\gamma \in \Gamma_0(N)$

$$f(\gamma z) = j(\gamma z)^{2k} f(z),$$

together with the condition that f be holomorphic in the cusps of $\Gamma_0(N)$. The usual way to do this is to define the Fourier expansion of f in each cusp and require that no negative terms occur. This is easily done at $i\infty$, where the Fourier expansion must look like

$$f(z) = \sum_{n\geq 0} a(n)e(nz). \tag{1}$$

For other cusps and $2k$ odd this is a little bit trickier and is best done using a cover of $\mathrm{SL}(2,\mathbb{R})$ (see (Shimura, 1973) or (Koblitz, 1984)). For our purposes it is enough to impose the equivalent growth condition on the invariant function $y^{k/2}|f(z)| = F(z)$ that

$$F(z) \ll y^A + y^{-A} \quad \text{for some} \quad A \geq 0 \tag{2}$$

and all $z \in \mathcal{H}$. Let $M_k(N)$ denote the space of all such functions; it is known to be finite dimensional. The subspace of cusp forms $S_k(N)$ consists of these $F \in M_k(N)$ whose zeroth Fourier coefficient in every cusp vanishes. For $k > 0$ this is equivalent to having (2) with $A = 0$.

The proof of Theorem A relies heavily on non-trivial estimates for the Fourier coefficients of cusp forms. This turns out to be rather harder when $2k$ is odd, which is the case needed. Let us recall the trivial bound of Hecke for a cusp form f and any k:

$$|a(n)| \underset{f}{\ll} n^{k/2}. \tag{3}$$

The proof is easy. For any $y > 0$

$$a(n)e^{-2\pi ny} = \int_0^1 e(-nx)f(x+iy)\,dx$$

and so using (2) with $A = 0$ gives

$$\begin{aligned} |a(n)| &\leq e^{2\pi ny}y^{-k/2}\int_0^1 F(x+iy)\,dx \\ &\ll e^{2\pi ny}y^{-k/2}. \end{aligned}$$

Taking $y = 1/n$ gives (3).

Hecke's bound certainly can fail for non-cusp forms; consider the easiest example when weight $k = 4$ of the Eisenstein series

$$E_4(z) = c_4 \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} (cz+d)^{-4} = 1 + 240\sum_{n=1}^{\infty}\sigma_3(n)e(nz), \tag{4}$$

which has $\sigma_3(n) = \sum_{d|n} d^3$ and cannot be bounded by a constant times n^2.

It is an important fact that one can make enough modular forms via Eisenstein series to subtract off the growth of an arbitrary modular form in the cusps, leaving a cusp form. This is harder for $k = \frac{1}{2}, 1, \frac{3}{2}$, and 2 since then the Eisenstein series do not converge absolutely. In fact, in these cases one is stuck dealing with non-holomorphic modular forms. This turns out to be the main difference between Theorems A and B.

4. Theta Series with Harmonic Polynomials

The relevance of modular forms to the Linnik problems is through the concept of a *Weyl sum*. Recall that for a finite set of points X_n on $S^1 = \mathbb{R}/\mathbb{Z}$, the Weyl criterion for equidistribution of X_n with respect to Lebesgue measure as $n \to \infty$ is that for each $m \in \mathbb{Z}$, $m \neq 0$,

$$\frac{1}{\#X_n} \sum_{\theta \in X_n} e(n\theta) \to 0$$

as $n \to \infty$. The situation for our points Ω_n on S^2 is very similar. Observe that

$$\left(\frac{x+iy}{|x+iy|}\right)^m = e(m\theta)$$

and

$$\left(\frac{x-iy}{|x-iy|}\right)^m = e(-m\theta)$$

if $\theta = \arg(x+iy)/2\pi$, for $m > 0$. Now $(x+iy)^m$ and $(x-iy)^m$ are homogeneous harmonic polynomials on $\mathbb{R}^2$. This example generalizes beautifully to $\mathbb{R}^n$. In particular for $\mathbb{R}^3$ it can be shown that any $f \in C^2(S^2)$ can be uniformly approximated by a finite sum of homogeneous harmonic polynomials in $\mathbb{R}^3$ restricted to S^2 (for a proof see Stein (Stein and Weiss, 1971, Corollary 2.3, p. 141)). Thus the Weyl criteria for uniform distribution of the lattice points on S^2 requires that we prove that for any homogeneous harmonic polynomial $P(x)$ on $\mathbb{R}^3$ of degree $\ell > 0$

$$\frac{1}{\#\Omega_n} \sum_{X \in \Omega_n} P(X) \to 0 \quad \text{as} \quad n \to \infty,$$

as in Theorem A. Equivalently, we require

$$\sum_{\substack{\alpha \in \mathbb{Z}^3 \\ |\alpha|^2 = n}} P\left(\frac{\alpha}{|\alpha|}\right) = o(r_3(n))$$

where $r_3(n) = \#\{\alpha \in \mathbb{Z}^3 : |\alpha|^2 = n\}$.

PROPOSITION 4.1. *The theta series*

$$\theta(z, P) = \sum_{\alpha \in \mathbb{Z}^3} P(\alpha) e(|\alpha|^2 z) = \sum_{n} r(n; P) e(nz)$$

is a holomorphic modular form of weight $\frac{3}{2} + \ell$ *for* $\Gamma_0(4)$*, which is a cusp form if* $\ell > 0$. *Also,* $\theta(z, P) = 0$ *unless* ℓ *is even.*

Proof. See (Shimura, 1973). □

When $\ell = \deg P = 0$ we have

$$\theta(z,1) = \theta^3(z) = \sum_{n\geq 0} r_3(n)e(nz).$$

To prove Theorem A, we need two ingredients:

(L) $r_3(n) \gg_\varepsilon n^{1/2-\varepsilon}$ for n as in Theorem A and all $\varepsilon > 0$,

(U) $|r(n,P)| \ll n^{k/2-1/4-\delta}$ for n square-free and some fixed $\delta > 0$, when $\ell > 0$.

To see this, note $\sum_{|\alpha|^2=n} P(\alpha/|\alpha|) = n^{-\ell/2} r(n;P)$ and $k/2 - \frac{1}{4} = \ell/2 + \frac{1}{2}$, so (U) says equivalently $\sum_{|\alpha|^2=n} P(\alpha/|\alpha|) \ll n^{1/2-\delta} = o(r_3(n))$. As we shall review below, (L) follows from classical results of Gauss and Siegel, but with an ineffective constant.

5. Linnik Problem for Squares and the Shimura Lift

At this point we see how far from (U) Hecke's exponent $k/2$ is. Before turning to this problem in earnest, let us treat a related problem that leads to integral weights, namely the distribution of rational points on S^2. These points are in one-one correspondence with the primitive $(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{Z}^3$ with $|\alpha|^2 = m^2$ via $\alpha \mapsto (1/m)\alpha$, $m > 0$. Here m is the height of the point. This easily leads us to consider the Linnik problem on S^2 for $n = m^2$.

Building an earlier results of Stieltjes, Hurwitz showed that

$$\sum_{n=1}^{\infty} r_3(n^2)n^{-s} = 6(1-2^{1-s})\frac{\zeta(s)\zeta(s-1)}{L(s,\chi_{-4})},$$

where $\chi_{-4}(\cdot) = (\frac{-4}{\cdot})$ is the Kronecker symbol. One easily derives from this that for odd n

$$r_3(n^2) \gg n, \tag{5}$$

which is even better than (L). This phenomenon was generalized by Shimura and is called the Shimura lift. In our case we can infer for $\ell > 0$ that there is a cusp form $F(z) = \Sigma a(n)e(nz)$ of weight $2\ell + 2$ for $\Gamma_0(2)$ such that

$$\sum_{n=1}^{\infty} r(n^2,P)n^{-s} = \frac{\sum_1^\infty a(n)n^{-s}}{L(s-\ell,\chi_{-4})}.$$

(see (Niwa, 1975)).

Thus

$$r(n^2,P) = \sum_{d|n} a(d)\mu(\tfrac{n}{d})\chi_{-4}(\tfrac{n}{d})(\tfrac{n}{d})^\ell$$

and so in place of (U) we need a bound for $a(n)$ of the form

$$|a(n)| \ll n^{\ell+1-\delta} \quad \text{as} \quad k - \tfrac{1}{2} = \ell + 1,$$

in order to beat the lower bound $r_3(n^2) \gg n$. Thus any non-trivial bound *for weight* $2\ell + 2$ cusp forms gives Theorem A for squares. It is then an easy matter to restrict to primitive points and derive the uniform distribution of rational points of a given height on S^2 as the height tends to infinity.

6. Nontrivial Estimates for Fourier Coefficients

At first look, the methods we shall apply to establish non-trivial estimates for the Fourier coefficients of cusp forms of integral and half-integral weights appear to be the same. However, there is a striking difference. Roughly speaking, one must overcome the bound given by Weil's bound for Kloosterman sums in the half-integral weight case. In fact, this bound is more appropriately called "trivial," as we will see.

The story about obtaining non-trivial bounds in the integral weight case has a complex plot. Here I will describe the Kloosterman sums approach. It will be observed that the role of Hecke operators has been ignored so far. Such an omission becomes a serious liability in the integral weight case but, since their role in the half-integral weight case is less central, at least for our purposes here, we will continue to not emphasize them.

Historically speaking, the first approach to obtaining non-trivial estimates for Fourier coefficients was via the circle method. Kloosterman produced his sums in this context and by non-trivially estimating them solved an important problem on the representations of integers by positive definite integral quadratic forms in four variables. Later it was found by Petersson and Selberg that one could take direct advantage of automorphy by constructing Poincaré series.

Consider for $\Gamma = \Gamma_0(N)$ and $m \geq 0$ the function

$$P_m(z,k) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} j(\gamma, z)^{-2k} e(m\gamma z),$$

which converges absolutely and uniformly on compact subsets of $\mathcal{H}$, provided that $k > 2$. It is not hard to show that $P_m \in S_k(N)$ for $m > 0$ and in fact they span S_k. Consider that for $f \in S_k(N)$ with $f(z) = \sum_1^\infty a(n)e(nz)$

$$\begin{aligned}\langle P_m, f\rangle &= \int_{\Gamma\backslash\mathcal{H}} P_m(z)\bar{f}(z)y^k \frac{dx\,dy}{y^2} \\ &= \int_{\Gamma\backslash\mathcal{H}} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} (j(\gamma,z))^{-2k} e(m\gamma z)\bar{f}(z)y^{k-2}\,dx\,dy\end{aligned}$$

$$\begin{aligned} &= \int_0^\infty \int_0^1 e(mz)\bar{f}(z)y^{k-2}\,dx\,dy \\ &= \bar{a}(m)(4\pi m)^{-k+1}\Gamma(k-1). \end{aligned} \tag{6}$$

Thus $\langle P_m, f\rangle = 0$ for all $m \geq 1 \Rightarrow a(m) = 0 \Rightarrow f \equiv 0$. It follows that it is enough for us to estimate the Fourier coefficients of P_m.

A nice calculation (see (Sarnak, 1990) or (Iwaniec, 1997)) shows that if we write

$$P_m(z,k) = \sum_{n>0} \hat{P}_m(n)e(nz)$$

then

$$\hat{P}_m(n) = 2(n/m)^{(k-1)/2}\left\{\delta_{m,n} + 2\pi i^{-k}\sum_{\substack{c\equiv 0(N)\\ c>0}} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right)K(m,n;c)c^{-1}\right\} \tag{7}$$

where

$$J_{k-1}(z) = \sum_{\ell\geq 0}\frac{(-1)^\ell}{\ell!\Gamma(\ell+k)}\left(\frac{z}{2}\right)^{k-1+2\ell}$$

is the J-Bessel function and

$$K(m,n;c) = \sum_{\substack{d(\mathrm{mod}\, c)\\ (d,c)=1}}\left(\frac{c}{d}\right)^{2k}\bar{\varepsilon}_d^{2k} e\left(\frac{m\bar{d}+nd}{c}\right)$$

is a Kloosterman sum of weight k, a kind of finite analogue of J_{k-1} (via an integral representation).

6.1. SUPPOSE $K > 2$ IS EVEN

This is enough to handle the Linnik problem for squares above, since the cusp form there was weight $2\ell + 2$. We shall ignore the case of odd integral k, even though in general this is very interesting.

For even k the Kloosterman sum is

$$K(m,n;c) = \sum_{\substack{d(c)\\ (d,c)=1}} e\left(\frac{m\bar{d}+nd}{c}\right)$$

and satisfies the famous Weil bound

$$|K(m,n;c)| \ll_\varepsilon c^{1/2+\varepsilon}.$$

The J-Bessel function satisfies for $x > 0$

$$J_{k-1}(x) \ll \min\left\{x^{k-1}, \frac{1}{\sqrt{x}}\right\}$$

so we may conclude from (7) that

$$\begin{aligned}\hat{P}_m(n) \ll_\varepsilon\ & n^{(k-1)/2}\left(\sum_{c\geq 4\pi\sqrt{mn}}\left(\frac{\sqrt{n}}{c}\right)^{k-1} c^{-1/2+\varepsilon}\right) \\ & +n^{(k-1)/2}\left(\sum_{c<4\pi\sqrt{mn}}\left(\frac{c}{\sqrt{n}}\right)^{1/2} c^{-1/2+\varepsilon}\right) \\ \ll\ & n^{k-1}\sum_{c\geq 4\pi\sqrt{mn}} c^{1/2-k+\varepsilon} + n^{k/2-3/4}\sum_{c<4\pi\sqrt{mn}} c^{\varepsilon} \\ \ll\ & n^{k/2-1/4+\varepsilon}.\end{aligned}$$

This gives the following result, hence the Linnik problem A *for squares*.

PROPOSITION 6.1. *For $k > 2$ even and $f \in S_k(N)$ with $f = \Sigma a(n)e(nz)$ we have*

$$a(n) \underset{\varepsilon}{\ll} n^{k/2-1/4+\varepsilon}.$$

REMARKS.

1. Any non-trivial bound for $K(m, n; c)$ yields a non-trivial bound for $a(n)$. This is what Kloosterman accomplished.
2. Another way to obtain a non-trivial estimate is the Rankin–Selberg method. This works well for integral weights but falls short of Weil's bound.

For general $k > 2$ and $f_j = \Sigma a_j(n)e(nz)$ an ortho-normal basis for $S_k(N)$, (6) gives

$$\begin{aligned}P_m(z, k) &= \sum_{j=1}^{J}\langle P_m, f_j\rangle f_j \\ &= \frac{\Gamma(k-1)}{(4\pi m)^{k-1}}\sum_{j=1}^{J}\bar{a}_j(m)f_j(z)\end{aligned}$$

and so

$$\hat{P}_m(n, k) = \frac{\Gamma(k-1)}{(4\pi m)^{k-1}}\sum_{j=1}^{J}\bar{a}_j(m)a_j(n).$$

Writing (7) for $\hat{P}_m(n,k)$ yields the Petersson formula. It is especially useful for estimations when $n = m$:

$$\frac{\Gamma(k-1)}{(4\pi n)^{k-1}} \sum_{j=1}^{J} |a_j(n)|^2 = 1 + 2\pi i^{-k} \sum_{c \equiv 0(N)} c^{-1} J_{k-1}\Big(\frac{4\pi n}{c}\Big) K(n,n;c). \qquad (8)$$

It is easily checked that this yields the estimate of Proposition 6.1 again.

For integral k this method reaches its limit here. One must introduce Hecke operators and interpret the Fourier coefficients of Hecke eigenforms algebraically. This led to Deligne's proof of the Ramanujan conjecture (Eichler proved the case $k = 2$).

THEOREM 6.2 (Deligne). *For $k \in \mathbb{Z}^+$ and $f \in S_k(N)$ we have*

$$a(n) \ll n^{(k-1)/2+\varepsilon}.$$

7. Salié Sums

When $2k$ is odd the Kloosterman sum still satisfies

$$|K(m,n;c)| \ll_{\varepsilon} c^{1/2+\varepsilon}$$

and Proposition 6.1 still holds, but now it is insufficient to get the Linnik problem since we needed to obtain for $\delta > 0$ and $k \geq 5/2$ the bound

$$|r(n,P)| \ll n^{k/2-1/4-\delta}. \qquad \text{(U)}$$

Perhaps it is appropriate that the exponent $k/2 - \frac{1}{4}$ is in fact "trivial" in the sense that Weil's bound in this case is entirely elementary. This is due to the fact that the Kloosterman sum can be evaluated, a fact observed in special cases by Salié. This evaluation is one of the keys behind Iwaniec's result. In this section we give a recent proof of Salié's result found by Árpád Tóth via Gauss sums (Tóth, 2005).

For the case $k = \frac{3}{2} + \ell$, ℓ even, the Kloosterman sum is

$$K(m,n;c) = \sum_{d(c)} \varepsilon_d \Big(\frac{c}{d}\Big) e\Big(\frac{md + n\bar{d}}{c}\Big).$$

This sum can be evaluated in a simpler form. By (8) we only need the case $m = n$.

An application of the Chinese reminder theorem and quadratic reciprocity puts the main behaviour on the Salié sum for $q > 0$ odd (factor out the even part)

$$S(m, n; q) = \sum_{a (\bmod q)} \left(\frac{a}{q}\right) e\left(\frac{ma + n\bar{a}}{q}\right).$$

The Jacobi symbol makes the Salié sum a finite analogue of J_{k-1} for $2k$ odd, which is elementary; for instance,

$$J_{1/2}(z) = \sqrt{\frac{2}{\pi z}} \sin z. \tag{9}$$

By changing variables when $(n, q) = 1$ we have

$$S(n, n, q) = \left(\frac{n}{q}\right) S(n^2, 1, q).$$

The analogue of (9) is

PROPOSITION 7.1.

$$S(n^2, 1, q) = \varepsilon_q \sqrt{q} \sum_{x^2 \equiv 1 (q)} e\left(\frac{2xn}{q}\right).$$

Tóth's proof. We use the Gauss sum

$$G(a, b; q) = \sum_{x(q)} e\left(\frac{ax^2 + bx}{q}\right)$$

with evaluation

$$G(a, 0; q) = \varepsilon_q \sqrt{q} \left(\frac{a}{q}\right).$$

Now let $A = \sum_{x^2 \equiv n^2 (q)} e(2x/q)$ so that we must show

$$S(n^2, 1; q) = \varepsilon_q \sqrt{q} A.$$

Now

$$\begin{aligned} A &= \frac{1}{q} \sum_{x(q)} e\left(\frac{2x}{q}\right) \sum_{a(q)} e\left(\frac{a(x^2 - n^2)}{q}\right) \\ &= \frac{1}{q} \sum_{a(q)} G(a, 2; q)\, e\left(\frac{-an^2}{q}\right) \\ &= \frac{1}{q} \sum_{(a,q)=1} G(a, 2; q)\, e\left(\frac{-an^2}{q}\right) \end{aligned}$$

since $G(a,b;q) = 0$ if $(a,q) \nmid 2$ and q is odd (exercise). But for $(a,q) = 1$

$$\begin{aligned} G(a,2;q) &= e\left(\frac{-\bar{a}}{q}\right)G(a,0;q) \\ &= e\left(\frac{-\bar{a}}{q}\right)\left(\frac{a}{c}\right)\varepsilon_q\sqrt{q} \end{aligned}$$

so

$$A = \frac{\varepsilon_q}{\sqrt{q}}\sum\left(\frac{a}{c}\right)e\left(\frac{-an^2-\bar{a}}{q}\right),$$

which gives the result since $A = \bar{A}$. □

8. An Estimate of Iwaniec

In 1987 Iwaniec (Iwaniec, 1987) proved

THEOREM 8.1. *Let $f \in S_k(N)$ with $2k \geq 5$ odd. Then, for n square-free,*

$$|a(n)| \underset{\varepsilon}{\ll} n^{k/2-1/4-1/28+\varepsilon}.$$

REMARK. By the Shimura lift this holds for all n. It also holds for forms with $k = \frac{1}{2}, \frac{3}{2}$ but now the square-free condition is needed.

Iwaniec's estimate makes use of an equivalent form of Proposition 7.1, namely that for q odd and $(n,q) = 1$

$$S(n,n;q) = \left(\frac{n}{q}\right)\varepsilon_q\sqrt{q}\sum_{\substack{ab=q\\(a,b)=1}} e\left(2n\left(\frac{\bar{a}}{b}-\frac{\bar{b}}{a}\right)\right). \tag{10}$$

He uses a lovely embedding idea in conjunction with the Petersson formula; cusp forms for $\Gamma_0(N)$ are also cusp forms for $\Gamma_0(M)$, if $N|M$. This, together with positivity leads to

$$(\log P)^{-1}\frac{|a(n)|^2}{n^{k-1}} \ll \frac{P}{\log P} + \sum_{P\leq p\leq 2P}\left|\sum_{c\equiv(\mathrm{mod}\,pN)}\frac{K(n,n;c)}{c}J_{k-1}\left(\frac{4\pi n}{c}\right)\right|. \tag{11}$$

By exploiting the bilinear form of (10) he was able to give eventually the bound

PROPOSITION 8.2.

$$\sum_{P\le p\le 2P} |K_{N_P}(x)| \underset{\varepsilon}{\ll} [xP^{-1/2}+xn^{-1/2}+(x+n)^{5/8}(x^{1/4}P^{3/8}+n^{1/8}x^{1/8}P^{1/4})](xnp)^{\varepsilon},$$

where

$$K_Q(x) = \sum_{\substack{c\le x\\ c\equiv 0(Q)}} c^{-1/2}K(n,n;c)e\left(\frac{2nv}{c}\right) \quad \textit{for } v = 0, 1, -1.$$

When combined with (11), this eventually leads to Theorem 8.1. Of course, this brief description hardly does justice to Iwaniec's argument. Sarnak has given an excellent treatment of the essential ideas in (Sarnak, 1990) and for full details the best reference is Iwaniec's original paper. Iwaniec later gave a different and in some ways simpler proof of theorem (with a weaker exponent) in (Iwaniec, 1997).

9. Theorems of Gauss and Siegel

In order to complete the proof of Theorem A we must now prove (L), since (U) follows from Iwaniec's estimate with any $\delta < 1/28$. In the Disquisitiones, Gauss proved that $r_3(n)$ is related to a class number. Suppose n is square-free. Then for $d = \operatorname{disc}\mathbb{Q}(\sqrt{-n})$ Gauss's formula can be put in the simple form

$$r_3(n) = 12H(d)\left(1 - \left(\frac{d}{2}\right)\right),$$

where $(d/2)$ is the Kronecker symbol. But Siegel proved (see (Iwaniec and Kowalski, 2004)) that

$$H(d) \underset{\varepsilon}{\gg} |d|^{1/2-\varepsilon}$$

for any $\varepsilon > 0$, but with an ineffective constant. Nonetheless, this gives (L), but it should be observed that we are forced to obtain (U) with a power savings—nothing less suffices. On the other hand, *any* $\delta > 0$ is enough.

The proof of Siegel's theorem is based on the class number formula of Dirichlet. Consider the Eisenstein series for $\Gamma = \mathrm{SL}(2,\mathbb{Z})$

$$E(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} (\operatorname{Im}\gamma z)^s, \quad \operatorname{Re}s > 1,$$

for which $\zeta(2s)E(z,s)$ has an analytic continuation with a simple pole at $s = 1$. Now

$$\zeta(2s)\sum_{z_Q\in\Lambda_d} E(z_Q,s) = \left(\frac{|d|}{4}\right)^{s/2} L(s,\chi_d)\zeta(s), \tag{12}$$

and taking residues at $s = 1$ gives the class number formula

$$H(d) = c|d|^{1/2}L(1,\chi_d).$$

Siegel's theorem is based on properties of $L(1,\chi_d)$.

10. The Nonholomorphic Case (Duke, 1988)

The proof of Theorem B follows along similar lines as the proof of Theorem A, but now we are forced to consider non-holomorphic modular forms. The first step, the identification of the Weyl sums, is accomplished via the spectral decomposition of the hyperbolic Laplacian on $\pm\Gamma\backslash\mathcal{H}$ for $\Gamma = \mathrm{SL}(2,\mathbb{Z})$.

We are lead naturally to consider the following two types of sums

(E) $\sum_{z_Q\in\Lambda_d} E(z_Q, s)$ for $\mathrm{Re}s = \frac{1}{2}$.

(C) $\sum_{z_Q\in\Lambda_d} \varphi(z_Q)$ for φ a Maass cusp form with $\Delta\varphi = \lambda\varphi$, $\Delta = -y^2(\partial_x^2 + \partial_y^2)$.

We just saw that (E) at $s = 1$ leads to the class number formula which indeed gives the same lower bound via Siegel's theorem that we must overcome. On $\mathrm{Re}s = \frac{1}{2}$ the problem becomes by (12) to estimate in terms of $|d|$ for some $\delta > 0$

$$L(\tfrac{1}{2} + it,\chi_d) \ll |d|^{1/4-\delta}. \tag{13}$$

This is precisely what Burgess (Burgess, 1963) accomplished in 1963, when he applied the RH for curves to get any $\delta < \frac{1}{16}$. Note that we also use the estimate $|\zeta(1 + 2it)| \gg \log(|t| + 2)^{-1}$ of de la Vallée Poussin.

To treat (C), we must generalize the theta function construction of Theorem A. This entails using a theta series for indefinite ternary forms, originally constructed by Siegel. A "theta lift" found in this context by Maass allows one to write (C) in terms of the dth Fourier coefficient of a Maass cusp form of weight $\frac{1}{2}$. An important refinement of the Maass construction was given by Katok and Sarnak (Katok and Sarnak, 1993) that identifies explicitly the eigenvalue dependence.

Although this is technically quite involved, conceptually it is not much different than the holomorphic case. One must replace the Petersson formula with a Kuznetsov formula that relates sums of Kloorterman sums to the whole (weight $\frac{1}{2}$) spectrum. This leads, with an appropriate choice of test functions, to the needed general version of Iwaniec's estimate (Proposition 8.2).

The Linnik problem for closed geodesics on $\pm\Gamma\backslash\mathcal{H}$ mentioned before is proven at the same time since the needed Weyl integrals occur as the dth coefficients of the same half-integral weight form, where now $d > 0$. One starts as before by considering the role of the Eisenstein series in the Dirichlet class number formula for real quadratic fields.

11. Transition to Subconvexity Bounds for L-Functions

The appearance of Burgess's bound (13) strongly hints that the problem of estimating non-trivially the Fourier coefficients of $\frac{1}{2}$-integral weight forms can be converted to the problem of bounding L-functions on the critical line. This is the case, with the paradigm being provided by Waldspurger's theorem. It turns out that in order to obtain non-trivial estimates in this way one must go beyond the convexity estimate of the Phragmen–Lindelöf Theorem, hence the name subconvexity bounds (see (Iwaniec and Sarnak, 2000)). This has led to a number of recent developments in the analytic theory of L-functions, which is currently an extremely active area.

After a series of papers by Duke, Friedlander and Iwaniec on GL(2) L-functions (see (Duke et al., 2002) for references), various convolution L-functions have been considered with associated equidistribution problems. For subconvexity estimates other important new contributions have been made by, among others, Bernstein, Blomer, Conrey, Harcos, Kowalski, Liu, Michel, Reznikov, Sarnak, Vanderkam, Venkatesh, Ye, (see e.g. (Michel, 2004)) for some recent references). The mixture of ergodic methods with topics around subconvexity is an exciting new direction being pursued by Lindenstrauss and Venkatesh.

12. An Application to Traces of Singular Moduli

I will end by describing a recent application of Theorem B to the asymptotics of traces of singular moduli (Duke, 2006).

Recall the classical j-function on $\mathcal{H}$

$$j(z) = \frac{(1 + 240\sum_{n=1}^{\infty}\sigma_3(n)q^n)^3}{q\,\Pi_{n=1}^{\infty}(1-q^n)^{24}} = q^{-1} + 744 + 196884q + \ldots$$

where $q = e(z) = e^{2niz}$. Now $j(\gamma z) = j(z)$ for $\gamma \in \Gamma$ and $j(z) = j(E)$ is the j-invariant of the elliptic curve $E/\mathbb{C}$ determinant by $\mathbb{C}/L$, where $L = \{m + nz; m, n \in \mathbb{Z}\}$. For a negative discriminant d a point $z_Q \in \Lambda_d$ is called a CM point since $j(z_Q)$ is the j-invariant of the elliptic curve E which has CM by the order $\mathbb{Z}[z_d]$. In fact, all such curves occur this way. The values $j(z_Q)$ are called singular moduli and are known to be conjugate algebraic integers for $z_Q \in \Lambda_d$. Let $K = Q(\sqrt{d})$ have discriminant $-D$. The field: $K(j(z_d))$ is Abelian over K and unramified outside of (m) where $d = -Dm^2$, called a ring class field. If $d = -D$ is fundamental then $K(j(z_d))$ is the Hilbert class field of K, that is the maximal unramified Abelian extension of K whose degree is the class number $h(d)$ of K (see (Cox, 1989)). Let us restrict to the case of fundamental d. Here is a table of the first few values of $j(z_d)$ (see Table I).

TABLE I.

d	$j(z_d)$
-3	0
-4	12^3
-7	-15^3
-8	20^3
-11	-32^3
-15	$\frac{1}{2}(-191025 - 85995\sqrt{5})$, the first irrational value

Consider $\mathrm{Tr}\,(j(z_d)) = \sum_{\Lambda_d} j(z_Q)$, which for $d < -4$ fundamental is the sum of the conjugates of $j(z_d)$. Clearly $\mathrm{Tr}\,(j(z_d)) \in \mathbb{Z}$. We shall apply Theorem B to get a precise asymptotic for $\mathrm{Tr}\,(j(z_d))$. A crude asymptotic is

$$\mathrm{Tr}\,j(z_d) = (-1)^d e^{\pi\sqrt{|d|}} + O(e^{\alpha\pi\sqrt{|d|}})$$

for any fixed $\alpha > \frac{1}{2}$. This comes from an easy examination of the height of the other z_Q's in the sum and an estimation of their number. To state a much more refined result, consider the exponential sum for $c > 0$ (later alias–Salié sum)

$$S_d(c) = \sum_{x^2 \equiv d(c)} e(2x/c).$$

Note that $\frac{1}{2}S_d(4) = (-1)^d$. The refinement is

COROLLARY. *As $d \to -\infty$ through fundamental discriminants*

$$\frac{1}{h(d)}\left(\mathrm{Tr}\,j(z_d) - \frac{1}{2}\sum_{\substack{0<c<2\sqrt{d}\\ c\equiv 0(4)}} S_d(c)e^{4\pi\sqrt{|d|}/c}\right) \to 720.$$

An equivalent form of this result was conjectured recently by Bruinier, Jenkins and Ono. It is remarkable that the constant 720 is an integer!

To see that this result is a consequence of Theorem B, fix $\varepsilon > 0$ and consider for a smooth (C^∞) test function $\psi\colon \mathbb{R}^+ \to [0,1]$ that is 0 on $[0,1]$ and 1 on $[1+\varepsilon, \infty)$, the Γ-invariant Poincaré series

$$h_\varepsilon(z) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \psi(\mathrm{Im}\,\gamma z)e(-\gamma z).$$

Here Γ_∞ consists of those $\gamma \in \Gamma$ that act as translations. Clearly for $\mathrm{Im}\,z > 1 + \varepsilon$ we have

$$h_\varepsilon(z) = e(-z)$$

and so $f(z) = j(z) - h_\varepsilon(z)$ is C^∞, Γ-invariant and bounded on $\mathcal{H}$. By Theorem B we have that as $d \to -\infty$

$$\frac{1}{h(d)}\left(\sum_{z\in\Lambda_d} j(z) - \sum_{z\in\Lambda_d} h_\varepsilon(z)\right) \to \int_{\mathcal{F}} j(z) - h_\varepsilon(z)\, d\mu.$$

Now,

$$\sum_{z\in\Lambda_d} h_\varepsilon(z) = \sum_{\operatorname{Im} z_Q>1} e(-z_Q) + O(\varepsilon h(-d)),$$

after applying again Theorem B to a suitable function.

Also,

$$\sum_{\operatorname{Im} z_Q>1} e(-z_Q) = \tfrac{1}{2} \sum_{\substack{0<c<2\sqrt{|d|}\\ c\equiv 0(4)}} S_d(c) e^{4\pi\sqrt{|d|}/c}$$

comes from the well known Gauss parametrization of roots of $x^2 \equiv d(c)$.

Next we need to evaluate

$$\int_{\Gamma\backslash\mathcal{H}} j(z) - h_\varepsilon(z)\, d\mu = \lim_{y\to\infty} \int_{\mathcal{F}_Y} j(z)\, d\mu$$

where $\mathcal{F}_Y = \{z \in \mathcal{F}; \operatorname{Im} z < Y\}$ since $\int_{\mathcal{F}_Y} h_\varepsilon(z)\, d\mu = 0$. Lerche, Schellenkens, and Warner showed how to evaluate such an integral using Stokes's theorem (see (Borcherds, 1998)). One uses the Eisenstein series of weight 2: $E_2(z) = 1 - 24\sum_1^\infty \sigma_1(n) q^n$ and its non-holomorphic modular version

$$\widetilde{E}_2(z) = E_2(z) - \frac{3}{\pi} y^{-1}.$$

Since

$$\partial\widetilde{E}_2/\partial\bar{z} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\left(\frac{-3}{\pi y}\right) = \frac{3i}{2\pi y^2}$$

and $d\bar{z}\, dz = 2i\, dx\, dy$ we get by Stokes's theorem

$$\begin{aligned} \frac{3}{\pi}\int_{\mathcal{F}_Y} j(z)\frac{dx\, dy}{y^2} &= \int_{-1/2+iY}^{1/2+iY} j(x+iY)\widetilde{E}_2(x+iY)\, dx \\ &= \text{constant term of } j\widetilde{E}_2(x+iY) \\ &= 744 - 2Y - \frac{3}{\pi}Y^{-1} \to 720, \text{ as } Y \to \infty. \end{aligned}$$

To see that $\int_{\mathcal{F}_Y} h_\varepsilon(z)\, d\mu = 0$, simply integrate the cut-off Poincaré series

$$h_{\varepsilon,Y}(z) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \psi_Y(\operatorname{Im}\gamma z) e(-\gamma z)$$

where $\psi_Y(y) = \begin{cases} \psi(y), & y \leq Y \\ 0, & y > Y \end{cases}$, which coincides with h_ε on $\mathcal{F}_Y$. Thus,

$$\int_{\mathcal{F}_Y} h_\varepsilon \, d\mu = \int_{\mathcal{F}} h_{\varepsilon,Y} \, d\mu = 0.$$

Acknowledgements

I wish to thank Zeév Rudnick and the referee for several corrections and suggestions. Research supported in part by NSF Grant DMS–0355564.

References

Borcherds, R. E. (1998) Automorphic forms with singularities on Grassmannians, *Invent. Math.* **132**, 491–562.

Burgess, D. A. (1963) On character sums and L-series. II, *Proc. London Math. Soc. (3)* **13**, 524–536.

Cohen, P. B. (2005) Hyperbolic equidistribution problems on Siegel 3-folds and Hilbert modular varieties, *Duke Math. J.* **129**, 87–127.

Cox, D. A. (1989) *Primes of the form* $x^2 + ny^2$, A Wiley-Interscience Publication, New York, John Wiley & Sons Inc.

Duke, W. (1988) Hyperbolic distribution problems and half-integral weight Maass forms, *Invent. Math.* **92**, 73–90.

Duke, W. (1997) Some old problems and new results about quadratic forms, *Notices Amer. Math. Soc.* **44**, 190–196.

Duke, W. (2006) Modular functions and the uniform distribution of CM points, *Math. Ann.* **334**, 241–252.

Duke, W., Friedlander, J. B., and Iwaniec, H. (2002) The subconvexity problem for Artin L-functions, *Invent. Math.* **149**, 489–577.

Duke, W. and Schulze-Pillot, R. (1990) Representation of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids, *Invent. Math.* **99**, 49–57.

Iwaniec, H. (1987) Fourier coefficients of modular forms of half-integral weight, *Invent. Math.* **87**, 385–401.

Iwaniec, H. (1997) *Topics in classical automorphic forms*, Vol. 17 of *Graduate Studies in Mathematics*, Providence, RI, American Mathematical Society.

Iwaniec, H. and Kowalski, E. (2004) *Analytic number theory*, Vol. 53 of *American Mathematical Society Colloquium Publications*, Providence, RI, American Mathematical Society.

Iwaniec, H. and Sarnak, P. (2000) Perspectives on the analytic theory of L-functions, *Geom. Funct. Anal.* pp. 705–741.

Katok, S. and Sarnak, P. (1993) Heegner points, cycles and Maass forms, *Israel J. Math.* **84**, 193–227.

Koblitz, N. (1984) *Introduction to elliptic curves and modular forms*, Vol. 97 of *Graduate Texts in Mathematics*, New York, Springer-Verlag.

Linnik, Y. V. (1968) *Ergodic properties of algebraic fields*, Translated from the Russian by M. S. Keane. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 45, Springer-Verlag New York Inc., New York.

Michel, P. (2004) The subconvexity problem for Rankin-Selberg L-functions and equidistribution of Heegner points, *Ann. of Math. (2)* **160**, 185–236.

Niwa, S. (1975) Modular forms of half integral weight and the integral of certain theta-functions, *Nagoya Math. J.* **56**, 147–161.

Sarnak, P. (1990) *Some applications of modular forms*, Vol. 99 of *Cambridge Tracts in Mathematics*, Cambridge, Cambridge University Press.

Shimura, G. (1973) On modular forms of half integral weight, *Ann. of Math. (2)* **97**, 440–481.

Stein, E. M. and Weiss, G. (1971) *Introduction to Fourier analysis on Euclidean spaces*, Princeton, N.J., Princeton University Press.

Tóth, Á. (2005) On the evaluation of Salié sums, *Proc. Amer. Math. Soc.* **133**, 643–645 (electronic).

DISTRIBUTION MODULO ONE AND RATNER'S THEOREM

Jens Marklof
University of Bristol

1. Introduction

Measure rigidity is a branch of ergodic theory that has recently contributed to the solution of some fundamental problems in number theory and mathematical physics. Examples are proofs of quantitative versions of the Oppenheim conjecture (Eskin et al., 1998), related questions on the spacings between the values of quadratic forms (Eskin et al., 2005; Marklof, 2003; Marklof, 2002), a proof of quantum unique ergodicity for certain classes of hyperbolic surfaces (Lindenstrauss, 2006), and an approach to the Littlewood conjecture on the nonexistence of multiplicatively badly approximable numbers (Einsiedler et al., 2006).

In these lectures we discuss a few simple applications of one of the central results in measure rigidity: Ratner's theorem. We shall investigate the statistical properties of certain number theoretic sequences, specifically the fractional parts of $m\alpha$, $m = 1, 2, 3, \ldots$, (a classical, well understood problem) and of $\sqrt{m\alpha}$ (as recently studied in (Elkies and McMullen, 2004)). By exploiting equidistribution results on a certain homogeneous space $\Gamma\backslash G$, we will show that the statistical properties of these sequences can exhibit significant deviations from those of independent random variables. The "randomness" of other, more generic sequences such as $m^2\alpha$ and $2^m\alpha$ mod 1 has been studied extensively. We refer the interested reader to the review (Marklof, 2006), and recommend the papers (Rudnick and Sarnak, 1998; Rudnick and Zaharescu, 2002) as a first read.

These notes are based on lectures presented at the Institute Henri Poincaré Paris, June 2005, and at the summer school 'Equidistribution in number theory', CRM Montréal, July 2005. The author gratefully acknowledges support by an EPSRC Advanced Research Fellowship.

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 217–244.

2. Randomness of Point Sequences mod 1

Consider an infinite triangular array of numbers on the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ (which we represent as the unit interval $[0, 1)$ with its endpoints identified),

$$\begin{array}{llll} \xi_{11} & & & \\ \xi_{21} & \xi_{22} & & \\ \vdots & \vdots & \ddots & \\ \xi_{N1} & \xi_{N2} & \cdots & \xi_{NN} \\ \vdots & \vdots & & \ddots \end{array} \tag{1}$$

We assume that each row is ordered, i.e., $\xi_{Nj} \leq \xi_{N(j+1)}$, and are interested in quantifying statistical properties of the Nth row as $N \to \infty$. To simplify notation we will from now on drop the index N, and simply write ξ_j instead of ξ_{Nj}.

As we shall see later, many interesting statistical properties of a sequence on $\mathbb{T}$ can be derived from the knowledge of the number of elements in small subintervals of $\mathbb{T}$. Let χ denote the characteristic function of the interval $[-\frac{1}{2}, \frac{1}{2}) \subset \mathbb{R}$. That is, $\chi(x) = 1$ if $-\frac{1}{2} \leq x < \frac{1}{2}$ and $= 0$ otherwise. The characteristic function of the interval $[x_0 - \ell/2, x_0 + \ell/2) + \mathbb{Z} \subset \mathbb{T}$ $(\ell \leq 1)$ can be represented as

$$\chi_\ell(x) = \sum_{n \in \mathbb{Z}} \chi\left(\frac{x - x_0 + n}{\ell}\right). \tag{2}$$

The sum over n makes sure χ_ℓ is periodic. The number of elements ξ_j in the interval are therefore

$$S_N(\ell) = \sum_{j=1}^{N} \chi_\ell(\xi_j). \tag{3}$$

We will always assume that the rows in our triangular array become uniformly distributed mod 1. This means that for every x_0, ℓ,

$$\lim_{N \to \infty} \frac{1}{N} S_N(\ell) = \ell, \tag{4}$$

i.e., the proportion of elements in any given interval is asymptotic to the interval length ℓ.[1]

[1] If a sequence $\{\xi_j\}$ fails to be uniformly distributed but still has a resonable limiting density ρ, we may rescale the ξ_j to obtain a uniformly distributed sequence. This is done as follows. Suppose for every $x_0, \ell \lim_{N \to \infty} (1/N) S_N(\ell) = \int_{x_0-\ell/2}^{x_0+\ell/2} \rho(x)\, dx$, where the integrated density $N(x) = \int_0^x \rho(x')\, dx'$ is continuous and strictly increasing. We rescale the sequence $\{\xi_j\}$ by setting $\tilde{\xi}_j := N(\xi_j)$. Note that $N(\xi_j) \in [0, 1)$ for $\xi_j \in [0, 1)$. The new sequence $\{\tilde{\xi}_j\}$ is indeed uniformly distributed modulo one (exercise).

The aim is now to characterize the different degrees of "randomness" of the deterministic sequence $\{\xi_j\}$ in terms of their distribution in very small intervals with random center x_0. A convenient length scale is the average spacing between elements, which is $(1/N)$. We set

$$L = N\ell. \tag{5}$$

We assume x_0 is a random variable uniformly distributed on $\mathbb{T}$ with respect to Lebesgue measure dx_0. We will denote expectation values by

$$\langle \cdots \rangle = \int_0^1 \cdots dx_0. \tag{6}$$

It is easy to work out the expectation value for the number of elements in a random interval of size ℓ,

$$\langle S_N(\ell) \rangle = L. \tag{7}$$

The variance is much less trivial. Let us begin by deriving a convenient representation in terms of the pair correlation density. We have for the mean square (the "number variance")

$$\Sigma_N^2(L) := \langle [S_N(\ell) - L]^2 \rangle = \langle S_N(\ell)^2 \rangle - L^2 \tag{8}$$

and

$$\begin{aligned} \langle S_N(\ell)^2 \rangle &= \sum_{i,j=1}^{N} \sum_{m,n\in\mathbb{Z}} \int_0^1 \chi\left(\frac{\xi_i - x_0 + m}{\ell}\right) \chi\left(\frac{\xi_j - x_0 + n}{\ell}\right) dx_0 \\ &= \sum_{i,j=1}^{N} \sum_{m\in\mathbb{Z}} \int_{\mathbb{R}} \chi\left(\frac{\xi_i - x_0 + m}{\ell}\right) \chi\left(\frac{\xi_j - x_0}{\ell}\right) dx_0 \\ &= \ell \sum_{i,j=1}^{N} \sum_{m\in\mathbb{Z}} \Delta\left(\frac{\xi_i - \xi_j + m}{\ell}\right) \end{aligned} \tag{9}$$

where

$$\Delta(x) = \int_{\mathbb{R}} \chi(x - x_0)\chi(x_0)\, dx_0 = \max\{1 - |x|, 0\}. \tag{10}$$

Now the diagonal terms $i = j$ in the above double sum can be easily evaluated. We have

$$\ell \sum_{i=j=1}^{N} \sum_{m\in\mathbb{Z}} \Delta\left(\frac{m}{\ell}\right) = \ell\Delta(0) = \ell \tag{11}$$

for $\ell < 1$.

The *pair correlation function* (also called *two-point correlation function*) for the sequence $\{\xi_j\}$ is defined by

$$R_N^2(L,\psi) = \frac{1}{N}\sum_{i\neq j=1}^{N}\sum_{m\in\mathbb{Z}}\psi\left(\frac{\xi_i-\xi_j+m}{\ell}\right), \tag{12}$$

where ψ is taken from a class of sufficiently nice test functions (e.g. continuous with compact support such as Δ). With the above calculation we therefore have the identity

$$\Sigma_N^2(L) = L - L^2 + LR_N^2(L,\Delta). \tag{13}$$

This says that the asymptotic analysis of the pair correlation density will give us information on the number variance.

Note that by the Poisson summation formula

$$\sum_{m\in\mathbb{Z}} f(m) = \sum_{n\in\mathbb{Z}} \hat{f}(n), \tag{14}$$

where

$$\hat{f}(y) = \int_{\mathbb{R}} f(x)e(xy)\,dy, \qquad e(x) := \exp 2\pi\mathrm{i}x, \tag{15}$$

we have

$$R_N^2(L,\psi) = \frac{L}{N^2}\sum_{i\neq j=1}^{N}\sum_{n\in\mathbb{Z}}\hat{\psi}\left(\frac{Ln}{N}\right)e(n(\xi_i-\xi_j)). \tag{16}$$

Here ψ can be any function with absolutely convergent Fourier series (e.g. Δ).

2.1. DISTRIBUTION OF GAPS

A popular statistical measure is the distribution of gaps

$$s_j = N(\xi_{j+1}-\xi_j) \quad (j = 1,\ldots,N, \xi_{N+1} := \xi_1 + 1) \tag{17}$$

between consecutive elements (recall the ξ_j form an ordered sequence on $\mathbb{T}$). We have multiplied the actual gap $\xi_{j+1}-\xi_j$ by N, which means we are measuring spacings in units of the average gap $(1/N)$.

The gap distribution of the sequence $\xi_1,\ldots,\xi_N$ is defined as

$$P_N(s) = \frac{1}{N}\sum_{j=1}^{N}\delta(s-s_j) \tag{18}$$

where δ is a Dirac mass at the origin. The question we will investigate is whether $P_N(s)$ has a limiting distribution $P(s)$. That is, does there exist a probability density $P(s)$ such that for every bounded continuous function $g\colon \mathbb{R} \to \mathbb{R}$,

$$\lim_{N\to\infty} \int_0^\infty g(s) P_N(s)\, ds = \int_0^\infty g(s) P(s)\, ds. \tag{19}$$

The first question in convergence of probability measures is the problem of tightness.

LEMMA 2.1. *The sequence of probability measures* $\{P_N(s)\, ds\}$ *is tight on* $\mathbb{R}$. *That is, for every* $\varepsilon > 0$ *there is a* $K > 0$ *such that for all* N

$$\int_{|s|>K} P_N(s)\, ds < \varepsilon. \tag{20}$$

Proof. We have

$$\begin{aligned} \int_{|s|>K} P_N(s)\, ds &= \frac{1}{N} \#\{j \le N : s_j \ge K\} \\ &\le \frac{1}{N} \sum_{j=1}^{N} \frac{s_j}{K} \chi_{[K,\infty)}(s_j) \le \frac{1}{N} \sum_{j=1}^{N} \frac{s_j}{K} \\ &= \frac{1}{K} \sum_{j=1}^{N} (\xi_{j+1} - \xi_j) = \frac{1}{K}. \end{aligned} \tag{21}$$

Denote by $E_N(k, L)$ the probability of finding k elements in the randomly shifted interval $[x_0, x_0 + L/N)$, i.e.,

$$E_N(k, L) := \operatorname{meas}\{x_0 \in \mathbb{T} : S_N(\ell) = k\}. \tag{22}$$

The following theorem explains the relation between $P(s)$ and the probability $E(0, L)$.

THEOREM 2.2. *Given a probability density* $P(s)$, *the following statements are equivalent*:

(i) $P_N(s) \to_w P(s)$.

(ii) $\lim_{N\to\infty} E_N(0, L) = E(0, L)$ *for all* $L > 0$, *where* $E(0, L)$ *is defined by*

$$\frac{d^2 E(0, L)}{dL^2} = P(L), \quad \lim_{L\to 0} E(0, L) = 1, \quad \lim_{L\to\infty} \frac{dE(0, L)}{dL} = 0. \tag{23}$$

Proof. We have

$$E_N(0, L) = \operatorname{meas}\left\{x_0 \in \mathbb{T}^2 : \#\left\{j : \xi_j \in \left[x_0, x_0 + \frac{L}{N}\right) + \mathbb{Z}\right\} = 0\right\}$$

$$= \sum_{j=1}^{N} \operatorname{meas}\left\{x_0 \in [\xi_j, \xi_{j+1}) : \#\left\{j : \xi_j \in \left[x_0, x_0 + \frac{L}{N}\right) + \mathbb{Z}\right\} = 0\right\}$$

$$= \sum_{j=1}^{N} \left(\xi_{j+1} - \xi_j - \frac{L}{N}\right) \chi_{[L,\infty)}(N(\xi_{j+1} - \xi_j))$$

$$= \sum_{j=1}^{N} (\xi_{j+1} - \xi_j) - \sum_{j=1}^{N} (\xi_{j+1} - \xi_j) \chi_{[0,L)}(N(\xi_{j+1} - \xi_j))$$

$$- \frac{L}{N} \sum_{j=1}^{N} \chi_{[L,\infty)}(N(\xi_{j+1} - \xi_j)])$$

$$= 1 - \frac{1}{N} \sum_{j=1}^{N} g(s_j) \tag{24}$$

where

$$g(x) = \max\{0, x, L\} \tag{25}$$

is a bounded continuous function.

(i) ⇒ (ii). With the above choice of test function g, (i) implies

$$\lim_{N \to \infty} E_N(0, L) = F(L) := 1 - \int_0^L sP(s)\, ds - L \int_L^\infty P(s)\, ds. \tag{26}$$

Now

$$\frac{dF(L)}{dL} = -\int_L^\infty P(s)\, ds, \qquad \frac{d^2F(L)}{dL^2} = P(L), \tag{27}$$

and

$$\lim_{L \to 0} F(L) = 1, \qquad \lim_{L \to \infty} \frac{dF(L)}{dL} = 0. \tag{28}$$

(ii) ⇒ (i). Since the sequence of probability measures $P_N(s)$ is tight, it is relatively compact by the Helly–Prokhorov theorem (also often called Helly's theorem). That is, every subsequence of N contains a convergent subsequence N_i for which $P_{N_i}(s) \to_w P(s)$ as $i \to \infty$. This implies (recall the first part of the proof) that $E_{N_i}(0, L) \to E(0, L)$ for all $L > 0$. Hence every convergent subsequence has the limit $E(0, L)$, and thus every subsequence convergences.

2.2. INDEPENDENT RANDOM VARIABLES

In order to understand which statistical behaviour we should expect for the deterministic sequences we will study later, let us assume the vector $\xi =$

$(\xi_1, \ldots, \xi_N)$ is a uniformly distributed random vector on $\mathbb{T}^N$ with respect to Lebesgue measure $d\xi = d\xi_1, \ldots, d\xi_N$. (This means the ξ_j are independent uniformly distributed random variables.) We can ignore the issue of ordering the ξ_j here because of the symmetry of the measure dx under permutation of coordinates. Expectation values and associated probabilities of a random variable $X = X(\xi)$ will be defined as

$$\mathbb{E}X = \int_{\mathbb{T}^N} X \, d\xi, \tag{29}$$

$$\mathrm{Prob}(X > R) = \mathrm{meas}\{\xi \in \mathbb{T}^N : X > R\}. \tag{30}$$

THEOREM 2.3. *There is a constant $C > 0$ such that, for all $\varepsilon > 0$, N, L,*

$$\mathrm{Prob}(|R_N^2(L,\psi) - L\hat{\psi}(0)| > \varepsilon) \le C \frac{L}{\varepsilon^2 N}. \tag{31}$$

Proof. First of all, we have for the expectation (the $n = 0$ term in (16))

$$\mathbb{E}R_N^2(L,\psi) = \frac{L(N-1)}{N} \hat{\psi}(0) = L\hat{\psi}(0)(1 + O(N^{-1})). \tag{32}$$

Secondly, for the variance of $R_N^2(L,\psi)$,

$$\begin{aligned} &\mathbb{E}|R_N^2(L,\psi) - \mathbb{E}R_N^2(L,\psi)|^2 \\ &= \frac{L^2}{N^4} \sum_{\substack{i \ne j \\ i' \ne j'}} \sum_{n,n' \in \mathbb{Z}} \hat{\psi}\left(\frac{Ln}{N}\right) \hat{\psi}\left(\frac{Ln'}{N}\right) \mathbb{E}[e(n(\xi_i - \xi_j) - n'(\xi_{i'} - \xi_{j'}))] \end{aligned} \tag{33}$$

Now

$$\mathbb{E}[e(n(\xi_i - \xi_j) - n'(\xi_{i'} - \xi_{j'}))] = \begin{cases} 1 & \text{if } n = n',\ i = i',\ j = j' \\ & \text{or if } n = -n',\ i = j',\ j = i' \\ 0 & \text{otherwise.} \end{cases} \tag{34}$$

This implies that

$$\mathbb{E}|R_N^2(L,\psi) - \mathbb{E}R_N^2(L,\psi)|^2 = \frac{L^2}{N^4} O(N^3/L) = O\left(\frac{L}{N}\right). \tag{35}$$

The above theorem implies that for a "generic" choice of the triangular array (1), we have

$$R_N^2(L,\psi) = L\hat{\psi}(0) + o(1) \tag{36}$$

in the limit $N \to \infty$, $\ell = L/N \to 0$. This implies for the variance

$$\Sigma_N^2(L) = L + o(L) \tag{37}$$

almost surely in the above limit.

Using standard techniques from probability theory, one can extend these results on the variance to the full distribution of a generic realization of the random sequence in a small randomly shifted interval. There are two scaling regimes.

Regime I. (*Central limit theorem*). In the limit $L \to \infty$, $N \to \infty$, $\ell = L/N \to 0$ we have

$$\operatorname{meas}\left\{ x_0 \in \mathbb{T} : \frac{S_N(\ell) - L}{\sqrt{\Sigma_N^2(L)}} > R \right\} \to \frac{1}{\sqrt{2\pi}} \int_R^\infty \mathrm{e}^{-t^2/2}\, dt \tag{38}$$

almost surely.

Regime II. (*Poisson limit theorem*). For L fixed, $N \to \infty$, we have

$$E_N(k, L) \to \frac{L^k}{k!}\mathrm{e}^{-L}. \tag{39}$$

almost surely.

3. *mα* mod One

We will now consider the statistical properties of the sequence given by the fractional parts of $m\alpha$, $m = 1, 2, 3, \ldots$ for some α. This problem was studied by Berry–Tabor, Pandey et al., Bleher, Mazel–Sinai and Greenman using continued fractions (see (Marklof, 2000) for detailed references). In particular, it is a classical result that there are at most three distinct values for the gaps occurring in $m\alpha$ mod 1 which already indicates a rather non-generic behavior of the sequence, see e.g. (Slater, 1967).

Here we will use the approach introduced in (Marklof, 2000) that has the advantage of avoiding continued fractions and thus allowing higher-dimensional generalizations, such as the analysis of the distribution of linear forms modulo one. It is also very close to the work of Elkies and McMullen on $\sqrt{m}$ mod 1 which we will discuss in the next section.

We will be interested in the regime where $L = N\ell$ is fixed (Poisson scaling regime). The number (3) of elements in an interval of size ℓ and centered at x_0 is then

$$S_N(\ell) = \sum_{m=1}^{N} \sum_{n\in\mathbb{Z}} \chi\left(\frac{N}{L}(m\alpha + n - x_0)\right)$$

$$= \sum_{(m,n)\in\mathbb{Z}^2} \chi_{(0,1]}\left(\frac{m}{N}\right)\chi_{[-L/2,L/2]}(N(m\alpha + n - x_0))$$

$$= \sum_{(m,n)\in\mathbb{Z}^2} \psi\left((m, n - x_0)\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}\begin{pmatrix} N^{-1} & 0 \\ 0 & N \end{pmatrix}\right) \tag{40}$$

where χ_I denotes the characteristic function of the interval $I \subset \mathbb{R}$ and

$$\psi(x, y) = \chi_{(0,1]}(x)\chi_{[-L/2,L/2]}(y) \tag{41}$$

is the characteristic function of a rectangle.

Define the Lie Group G by the semidirect product $\mathrm{SL}(2,\mathbb{R}) \ltimes \mathbb{R}^2$ with multiplication law

$$(M, \xi)(M', \xi') = (MM', \xi M' + \xi'), \tag{42}$$

where ξ, $\xi' \in \mathbb{R}^2$ are viewed as row vectors. This group has the matrix representation

$$(M, \xi) \mapsto \begin{pmatrix} M & 0 \\ \xi & 1 \end{pmatrix} \in \mathrm{SL}(3, \mathbb{R}). \tag{43}$$

The function

$$F(M, \xi) = \sum_{m\in\mathbb{Z}^2} \psi(mM + \xi) \tag{44}$$

defines a function on G. Note that, with ψ as above, the sum in (44) is always finite, and hence F is a piecewise constant function. Furthermore,

$$S_N(\ell) = F(M, \xi) \tag{45}$$

for the special choice

$$M = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}\begin{pmatrix} N^{-1} & 0 \\ 0 & N \end{pmatrix}, \quad \xi = (0, -x_0)M. \tag{46}$$

The crucial observation is now that F is left-invariant under the discrete subgroup $\Gamma = \mathrm{SL}(2, \mathbb{Z}) \ltimes \mathbb{Z}^2$, and hence F may be viewed as a piecewise constant function on the homogeneous space $\Gamma\backslash G$.

PROPOSITION 3.1. *$F(\hat{\gamma} g) = F(g)$ for all $\hat{\gamma} \in \Gamma$.*

Proof. We have the decomposition

$$\hat{\gamma} = (\gamma, n) = (\gamma, 0)(1, n) \tag{47}$$

for some $\gamma \in \mathrm{SL}(2,\mathbb{Z})$, $n \in \mathbb{Z}^2$. It is therefore sufficient to check the statement for elements of the form $(\gamma, 0)$ and $(1, n)$ separately. We have

$$\begin{aligned} F((1,n)(M,\xi)) &= F(M, nM+\xi) \\ &= \sum_{m\in\mathbb{Z}^2} \psi((m+n)M+\xi) \\ &= \sum_{m\in\mathbb{Z}^2} \psi(mM+\xi) \\ &= F(M,\xi) \end{aligned} \tag{48}$$

which proves one case, and

$$\begin{aligned} F((\gamma,0)(M,\xi)) &= F(\gamma M,\xi) \\ &= \sum_{m\in\mathbb{Z}^2} \psi(m\gamma M+\xi) \\ &= \sum_{m\in\mathbb{Z}^2} \psi(mM+\xi) \\ &= F(M,\xi) \end{aligned} \tag{49}$$

since $\gamma\mathbb{Z}^2 = \mathbb{Z}^2$.

Alternatively, F may be expressed as

$$F(g) = \sum_{\hat{\gamma}\in\pi(\Gamma)\backslash\Gamma} \psi(\pi(\hat{\gamma}g)) \tag{50}$$

with the projection

$$\begin{aligned} \pi\colon G &\to \mathbb{R}^2 \\ (M,\xi) &\mapsto \xi. \end{aligned} \tag{51}$$

From (50) the invariance under Γ is directly evident.

3.1. GEOMETRY OF $\Gamma\backslash G$

The aim is to find a good coordinate system for G. Since parametrizing $\mathbb{R}^2$ is obvious, we need to mainly worry about $\mathrm{SL}(2,\mathbb{R})$. The Iwasawa decomposition of an element $M \in \mathrm{SL}(2,\mathbb{R})$ is

$$M = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v^{1/2} & 0 \\ 0 & v^{-1/2} \end{pmatrix} \begin{pmatrix} \cos(\phi/2) & \sin(\phi/2) \\ -\sin(\phi/2) & \cos(\phi/2) \end{pmatrix} \tag{52}$$

where $\tau = u + \mathrm{i}v \in \mathbb{H} := \{\tau \in \mathbb{C} : \operatorname{Im}\tau > 0\}$ (the complex upper halfplane) and $\phi \in [0, 4\pi)$. This yields a $1-1$ map $\mathrm{SL}(2,\mathbb{R}) \to \mathbb{H} \times [0,4\pi)$. Left-multiplication becomes now an action of $\mathrm{SL}(2,\mathbb{R})$ on $\mathbb{H} \times [0,4\pi)$ given by the formula

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \left(\frac{a\tau + b}{c\tau + d}, \phi - 2\arg(c\tau + d) \right) \tag{53}$$

(this can be checked by a straightforward calculation). The fractional linear transformation of the τ component defines an (orientation preserving) isometry with respect to the Riemannian line element

$$ds^2 = \frac{du^2 + dv^2}{v^2} \tag{54}$$

and the transformation property of ϕ is identical to the direction of a tangent vector at $\tau \in \mathbb{H}$. Thus the group $\mathrm{PSL}(2,\mathbb{R}) := \mathrm{SL}(2,\mathbb{R})/\{\pm 1\} \simeq \mathbb{H} \times [0, 2\pi)$ can be identified with the unit tangent bundle $\mathrm{T}^1\mathbb{H}$ of $\mathbb{H}$. Similarly, $\mathrm{SL}(2,\mathbb{Z})\backslash \mathrm{SL}(2,\mathbb{R}) \simeq \mathrm{PSL}(2,\mathbb{Z})\backslash \mathrm{PSL}(2,\mathbb{R})$ can be identified with the unit tangent bundle of the modular surface $\mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H}$. A fundamental domain $\mathcal{F}$ for the action of $\mathrm{SL}(2,\mathbb{Z})$ on $\mathbb{H}$ is shown in Figure 1. We have

$$\begin{aligned} \mathcal{F} \;=\; & \{\tau \in \mathbb{H} : |\tau| > 1, \, |\operatorname{Re}\tau| < \tfrac{1}{2}\} \\ & \cup \{\tau \in \mathbb{H} : |\tau| \geq 1, \, \operatorname{Re}\tau = -\tfrac{1}{2}\} \\ & \cup \{\tau \in \mathbb{H} : |\tau| = 1, \, -\tfrac{1}{2} \leq \operatorname{Re}\tau \leq 0\}. \end{aligned} \tag{55}$$

Note that the modular surface is not compact, there is one cusp at $\mathrm{i}\infty$. It has however finite measure with respect to the Riemannian volume $v^{-2} du\, dv$.

In order to understand the geometry of all of $\Gamma\backslash G$, write

$$g = (1, \xi)(M, 0) \tag{56}$$

which gives a particular parametrization in terms of $\mathbb{R}^2$ and $\mathrm{SL}(2,\mathbb{R})$. Since Γ contains the subgroup $1 \ltimes \mathbb{Z}^2$, ξ can be parametrized by $\mathbb{T}^2 = \mathbb{Z}^2\backslash\mathbb{R}^2 \simeq [0,1)^2$. This concludes our analysis: we have found a $1-1$ parametrization of G in terms of

$$\mathrm{T}^1(\,\mathrm{SL}(2,\mathbb{R})\backslash\mathbb{H}) \times \mathbb{T}^2 \simeq \mathcal{F} \times [0, 2\pi) \times [0,1)^2. \tag{57}$$

That is, $\Gamma\backslash G$ is a (non-trivial) bundle over $\mathrm{T}^1(\mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H})$ with fibre $\mathbb{T}^2$.

3.2. DYNAMICS ON $\Gamma\backslash G$

Consider the one-parameter subgroup $\Phi^{\mathbb{R}} := \{\Phi^t\}_{t\in\mathbb{R}}$ where

$$\Phi^t = \left(\begin{pmatrix} \mathrm{e}^{-t/2} & 0 \\ 0 & \mathrm{e}^{t/2} \end{pmatrix}, 0 \right). \tag{58}$$

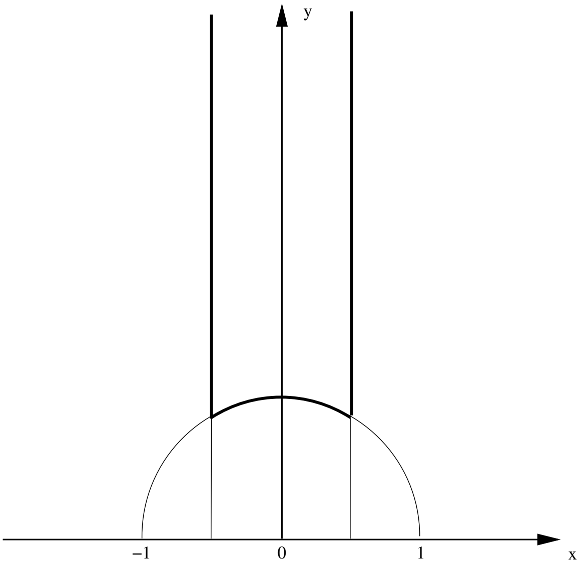


Figure 1. Fundamental domain of the modular group $\mathrm{SL}(2,\mathbb{Z})$ in the complex upper half plane

$\Phi^{\mathbb{R}}$ defines a flow on $\Gamma\backslash G$ by right multiplication,

$$\Gamma g \mapsto \Gamma g \Phi^t. \tag{59}$$

The remarkable observation is that our object of interest, $S_N(\ell)$, is related to a function F on $\Gamma\backslash G$ evaluated along an orbit of this flow:

$$S_N(\ell) = F(g_0 \Phi^t) \tag{60}$$

with $t = 2\log N$ and initial condition

$$g_0 = \left(\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}, (0, -x_0) \right). \tag{61}$$

Let us define

$$n_-(\alpha, y) = \left(\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}, (0, y) \right). \tag{62}$$

The subgroup $H = \{n_-(\alpha, y)\}_{(\alpha,y)\in\mathbb{R}^2}$ is Abelian and isomorphic to $\mathbb{R}^2$. Notice that

$$\Gamma \cap H = \{n_-(\alpha, y)\}_{(\alpha,y)\in\mathbb{Z}^2} \tag{63}$$

is a subgroup of H isomorphic to $\mathbb{Z}^2$. Therefore, for every fixed t, the set

$$\Gamma\backslash\Gamma H \Phi^t \tag{64}$$

describes a torus $\simeq \mathbb{T}^2$ embedded in $\Gamma\backslash G$; t parametrizes a continuous family of such tori.

We will now show that H parametrizes the unstable directions of the flow Φ^t. We employ the following parametrization of G. Write

$$g = n_-(\alpha, y)\Phi^s n_+(\beta, x), \tag{65}$$

where

$$n_+(\beta, x) = \left(\begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix}, (x, 0) \right). \tag{66}$$

We will write for short $g = (\alpha, y, s, \beta, x)$. The advantage of these coordinates is that the time evolution under Φ^t can be worked out very simply. We have the relation

$$(\alpha, y, s, \beta, x)\Phi^t = \Phi^t(\mathrm{e}^t\alpha, \mathrm{e}^{t/2}y, s, \mathrm{e}^{-t}\beta, \mathrm{e}^{-t/2}x). \tag{67}$$

Distances on $\Gamma\backslash G$ are measured by a left-G-invariant (since Γ acts on the left) Riemannian metric $d(g, g')$ on G. If $g = (\alpha, y, 0, 0, 0)$ and $g' = (\alpha', y', 0, 0, 0)$ are two initially close points, we have under the flow Φ^t (use the above formula and left-invariance of the metric)

$$\begin{aligned} d(g\Phi^t, g'\Phi^t) &= d((\mathrm{e}^t(\alpha - \alpha'), \mathrm{e}^{t/2}(y - y'), 0, 0, 0), (0, 0, 0, 0, 0)) \\ &\approx (\mathrm{e}^{2t}|\alpha - \alpha'|^2 + \mathrm{e}^t|y - y'|^2)^{1/2}. \end{aligned} \tag{68}$$

Hence (α, y) describe exponentially unstable directions of the flow, and by the same argument it is easy to see that (β, y) are the exponentially stable directions and s is of course the neutral flow direction. In particular we have the bound

$$d((\alpha, y, s, \beta, x)\Phi^t, (\alpha, y, 0, 0, 0)\Phi^t) = O(|s| + |\beta|\mathrm{e}^{-t} + |x|\mathrm{e}^{-t/2}) \tag{69}$$

for s, β, x bounded and $t > 0$. This follows directly from (67).

3.3. MIXING AND UNIFORM DISTRIBUTION

Recall that we are interested in the behaviour of the distribution of $S_N(\ell)$ for x_0 random and N large. At this point it will be convenient to also take α to be random, say, uniformly distributed in the interval $[a, b]$. We will see later that for fixed α there is no universal limiting distribution (an observation that is well known and related to the three gap theorem (Slater, 1967)).

We will use equidistribution on $\Gamma\backslash G$ to prove the following limit theorem, which asserts a limiting distribution different from Poissonian, cf. (39).

We will use the notation $\bar{g} = \Gamma g$.

THEOREM 3.2. *For any $L > 0$,*

$$\lim_{N\to\infty} \frac{1}{b-a} \operatorname{meas}\{(\alpha, x_0) \in [a,b] \times [0,1] : S_N(\ell) = k\} = E(k,L), \tag{70}$$

where

$$E(k,L) = \frac{1}{\mu(\Gamma\backslash G)} \mu(\bar{g} \in \Gamma\backslash G : F(g) = k). \tag{71}$$

Here F is the function defined in (44), and μ the Haar measure on G. An explicit formula for $d\mu$ in the Iwasawa coordinates is

$$d\mu = \frac{du\, dv}{v^2}\, d\phi\, dx\, dy. \tag{72}$$

It is possible to derive more explicit formulas for $E(k,L)$ from (71), but this requires some involved calculations which we will not pursue her. See (Strömbergsson and Venkatesh, 2005, Section 8), for details.

The key to the proof is the following equidistribution theorem.

THEOREM 3.3. *For any bounded, piecewise continuous*[2] $f\colon \Gamma\backslash G \to \mathbb{R}$

$$\lim_{t\to\infty} \frac{1}{b-a} \int_a^b \int_0^1 f(n_-(\alpha, y)\Phi^t)\, d\alpha\, dy = \frac{1}{\mu(\Gamma\backslash G)} \int_{\Gamma\backslash G} f\, d\mu. \tag{73}$$

Proof. It is well known that the flow Φ^t is mixing,[3] that is for any f, $h \in \mathrm{L}^2(\Gamma\backslash G)$

$$\lim_{t\to\infty} \int_{\Gamma\backslash G} f(g\Phi^t) h(g)\, d\mu = \frac{1}{\mu(\Gamma\backslash G)} \int_{\Gamma\backslash G} f\, d\mu \int_{\Gamma\backslash G} h\, d\mu. \tag{74}$$

Take f to be continuous an of compact support, and h the characteristic function of the set

$$S_\varepsilon = \Gamma\{(\alpha, y, s, \beta, y) : \alpha \in [a,b], y \in [0,1], s, \beta, x \in [-\varepsilon, \varepsilon]\}, \tag{75}$$

which forms an ε-neighbourhood of the embedded closed torus S_0. By the uniform continuity of f and (69), given any $\delta > 0$ there is an $\varepsilon > 0$ such that

$$\sup_{\substack{g\in S_\varepsilon \\ t>0}} |f(g\Phi^t) - f(n_-(\alpha, y)\Phi^t)| < \delta. \tag{76}$$

Haar measure in the local coordinates (α, y, s, β, y) reads (up to normalization)

$$d\mu = \mathrm{e}^{3s/2}\, ds\, d\alpha\, d\beta\, dx\, dy. \tag{77}$$

[2] i.e. the discontinuities are contained in a set of μ measure zero.

[3] This is guaranteed by a general theorem by Moore for semisimple Lie groups, which can be extended to the non-semisimple G considered here, cf. (Kleinbock, 1999).

We conclude that

$$\liminf_{t\to\infty} \frac{1}{b-a} \int_a^b \int_0^1 f(n_-(\alpha, y)\Phi^t)\, d\alpha\, dy = \frac{1}{\mu(\Gamma\backslash G)} \int_{\Gamma\backslash G} f\, d\mu + O(\delta) \quad (78)$$

and

$$\limsup_{t\to\infty} \frac{1}{b-a} \int_a^b \int_0^1 f(n_-(\alpha, y)\Phi^t)\, d\alpha\, dy = \frac{1}{\mu(\Gamma\backslash G)} \int_{\Gamma\backslash G} f\, d\mu + O(\delta), \quad (79)$$

where the implied constants are independent of ε. This works for any $\delta > 0$, and hence the limit must exist and equal $1/(\mu(\Gamma\backslash G)) \int_{\Gamma\backslash G} f\, d\mu$.

To extend the statement of the theorem to bounded continuous functions, we observe that it holds (trivially) for constant f, and therefore also for continuous functions f that are constant outside some compact set.

Let f be a bounded piecewise continuous function. Given any $\varepsilon > 0$ we can find continuous functions $f_\pm$, constant outside some constant set, such that

$$f_- \leq f \leq f_+ \quad (80)$$

and

$$\frac{1}{\mu(\Gamma\backslash G)} \int_{\Gamma\backslash G} (f_+ - f_-)\, d\mu < \varepsilon. \quad (81)$$

This implies

$$\begin{aligned} \liminf_{t\to\infty} \frac{1}{b-a} \int_a^b \int_0^1 & f(n_-(\alpha, y)\Phi^t)\, d\alpha\, dy \\ & \geq \liminf_{t\to\infty} \frac{1}{b-a} \int_a^b \int_0^1 f_-(n_-(\alpha, y)\Phi^t)\, d\alpha\, dy \\ & = \frac{1}{\mu(\Gamma\backslash G)} \int_{\Gamma\backslash G} f_- d\mu \\ & > \frac{1}{\mu(\Gamma\backslash G)} \int_{\Gamma\backslash G} f d\mu - 2\varepsilon. \end{aligned} \quad (82)$$

The analogous argument shows

$$\limsup_{t\to\infty} \frac{1}{b-a} \int_a^b \int_0^1 f(n_-(\alpha, y)\Phi^t)\, d\alpha\, dy < \frac{1}{\mu(\Gamma\backslash G)} \int_{\Gamma\backslash G} f d\mu + 2\varepsilon. \quad (83)$$

Taking $\varepsilon > 0$ arbitrarily small proves the theorem.

REMARK 3.4. An alternative proof of Theorem 3.3 follows from Ratner's theorem, since the subgroup $\{n_-(\alpha, y)\}_{\alpha, y\in\mathbb{R}}$ is generated by unipotent elements. We will get back to this later.

Proof. [Proof of Theorem 3.2] Apply Theorem 3.3 to the characteristic function of the set of $\bar{g} \in \Gamma\backslash G$ for which $F(g) = k$ (to make sure the characteristic function is piecewise continuous, check that the set has a boundary of μ measure zero).

REMARK 3.5. As we had mentioned earlier, there is no limiting distribution as in Theorem 3.2 if α is fixed, since there is no analog of the equidistribution result, Theorem 3.3. One can show, however, that if α is irrational we have for any continuous, compactly supported function

$$\int_{\mathbb{T}} f(n_-(\alpha, y)\Phi^t)\, dy = \bar{f}\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \mathrm{e}^{-t/2} & 0 \\ 0 & \mathrm{e}^{t/2} \end{pmatrix} + o(1) \quad (t \to \infty) \tag{84}$$

where $\bar{f}$ is a (non-constant!) continuous, compactly supported function on $\mathrm{SL}(2,\mathbb{Z})\backslash \mathrm{SL}(2,\mathbb{R})$ defined by

$$\bar{f}(M) = \int_{\mathbb{T}^2} f((1,\xi)(M,0))\, d\xi. \tag{85}$$

REMARK 3.6. If one however fixes $y = -x_0 \notin \mathbb{Q}$ and keeps α random, Ratner's theorem implies the following equidistribution result. For any bounded piecewise continuous $f\colon \Gamma\backslash G \to \mathbb{R}$

$$\lim_{t\to\infty} \int_{\mathbb{T}} f(n_-(\alpha, y)\Phi^t)\, d\alpha = \frac{1}{\mu(\Gamma\backslash G)} \int_{\Gamma\backslash G} f\, d\mu. \tag{86}$$

Hence the limiting distribution is universal (i.e. independent of y as long as y is irrational) and the same as for random y. Thus the probability of finding k points in the interval $x_0 - \ell/2, x_0 + \ell/2)$ with *fixed* center $x_0 \notin \mathbb{Q}$ has the limiting distribution

$$\lim_{N\to\infty} \mathrm{meas}\{\alpha \in \mathbb{T}^2 : S_N(\ell) = k\} = E(k, L), \tag{87}$$

the same as for *random* center. We will prove (3.3) in Section 5.

4. $\sqrt{m\alpha}$ mod One

The problem of the statistics of $\sqrt{m\alpha}$ mod 1 has been understood by Elkies and McMullen (Elkies and McMullen, 2004) in the case $\alpha = 1$ (and in principle also for all other rational α). The uniform distribution of $\sqrt{m\alpha}$ mod 1 may be shown by using the fact that $\sqrt{n+m} - \sqrt{n} \to 0$ for $n \to \infty$, m fixed (we leave this as an exercise). As in the last section, the key idea is the reduce the problem to equidistribution on a homogeneous space. Lucky for us, this homogeneous space will turn out to be $\Gamma\backslash G$ with the same G, Γ as encountered earlier.

We are as in the previous section interested in the "Poisson scaling limit", i.e. L is fixed. Now (we swap m and n in our notation)

$$S_N(\ell) = \sum_{n=1}^{N} \sum_{m\in\mathbb{Z}} \chi\left(\frac{N}{L}(\sqrt{n\alpha} - x_0 + m)\right). \tag{88}$$

The condition imposed on the summation can be re-written as

$$\left(x_0 - m - \frac{L}{2N}\right)^2 \leq n\alpha < \left(x_0 - m + \frac{L}{2N}\right)^2 \tag{89}$$

which amounts to

$$-\frac{L}{N}(x_0 - m) \leq n\alpha - (x_0 - m)^2 - \left(\frac{L}{2N}\right)^2 < \frac{L}{N}(x_0 - m). \tag{90}$$

Notice also that

$$|\sqrt{n\alpha} - (x_0 - m)| \leq \frac{L}{2N}. \tag{91}$$

This yields

$$\begin{aligned} S_N(\ell) = \sum_{(m,n)\in\mathbb{Z}^2} & \chi_{(0,1]}\left(\frac{x_0 - m + O(L/2N)}{\sqrt{N\alpha}}\right) \\ & \chi_{[-L,L)}\left(\frac{N^{1/2}[n\alpha - (x_0 - m)^2 - (L/2N)^2]}{N^{-1/2}(x_0 - m)}\right). \end{aligned} \tag{92}$$

A more convenient object would be

$$\widetilde{S}_{N,\varepsilon,\delta}(\ell) = \sum_{(m,n)\in\mathbb{Z}^2} \chi_{(-\varepsilon,1+\varepsilon]}\left(\frac{x_0 - m}{\sqrt{N\alpha}}\right)\chi_{[-L,L)}\left(\frac{N^{1/2}[n\alpha - (x_0 - m)^2] + \delta}{N^{-1/2}(x_0 - m)}\right). \tag{93}$$

For the right choices of ε (positive/negative) we obtain upper/lower bounds for $S_N(\ell)$ which would eventually allow us to infer the limiting distribution of $S_N(\ell)$ from $\widetilde{S}_{N,\varepsilon,\delta}(\ell)$ by taking $\delta \to 0$, $\varepsilon \to \pm 0$. We will ignore this technical point here and simply take

$$S_N(\ell) \approx \widetilde{S}_{N,0,0}(\ell) = \sum_{(m,n)\in\mathbb{Z}^2} \chi_{(0,1]}\left(\frac{x_0 - m}{\sqrt{N\alpha}}\right)\chi_{[-L,L)}\left(\frac{N^{1/2}[n\alpha - (x_0 - m)^2]}{N^{-1/2}(x_0 - m)}\right). \tag{94}$$

The manipulations we will now perform on the r.h.s. of (94) can be adapted step by step for more general values of $\delta, \varepsilon \neq 0$ (recommended exercise). We will use the shorthand $\widetilde{S}_N(\ell) := \widetilde{S}_{N,0,0}(\ell)$ in the following.

4.1. THE CASE $\alpha = 1$

We have, after substituting $(m, n) \to (-m, -n)$,

$$\widetilde{S}_N(\ell) = \sum_{(m,n)\in\mathbb{Z}^2} \chi_{(0,1]}\left(\frac{x_0 + m}{\sqrt{N}}\right)\chi_{(-L,L]}\left(\frac{N^{1/2}[n + (x_0 + m)^2]}{N^{-1/2}(x_0 + m)}\right), \tag{95}$$

an thus, after substituting $n \mapsto n + m^2$ in the sum over n,

$$\widetilde{S}_N(\ell) = \sum_{(m,n)\in\mathbb{Z}^2} \chi_{(0,1]}\left(\frac{x_0 + m}{\sqrt{N}}\right)\chi_{(-L,L]}\left(\frac{N^{1/2}(n + x_0^2 + 2mx_0)}{N^{-1/2}(x_0 + m)}\right). \tag{96}$$

We will now show that, in analogy with the previous section, we can find a function $F\colon \Gamma\backslash G \to \mathbb{R}$ of the form

$$F(M, \xi) = \sum_{m\in\mathbb{Z}^2} \psi(mM + \xi) \tag{97}$$

so that

$$\widetilde{S}_N(\ell) = F(g) \tag{98}$$

for a suitable choice of $g \in G$ and a piecewise continuous $\psi\colon \mathbb{R}^2 \to \mathbb{R}$ with compact support. To this end define

$$\psi(x, y) = \chi_{(0,1]}(x)\chi_{(-L,L]}\left(\frac{y}{x}\right) \tag{99}$$

(which indeed has compact support: it is the characteristic function of a triangle). Now consider the one parameter subgroup $\{n_1(x)\}_{x\in\mathbb{R}}$ with

$$n_1(x) = \left(\begin{pmatrix} 1 & 2x \\ 0 & 1 \end{pmatrix}, (x, x^2)\right) \tag{100}$$

(check that this indeed yields a one parameter group). Then the choice (set $t = \log N$, $x = x_0$)

$$\begin{aligned}(M, \xi) &= n_1(x_0)\Phi^t \\ &= \left(\begin{pmatrix} N^{-1/2} & 2x_0N^{1/2} \\ 0 & N^{1/2} \end{pmatrix}, (N^{-1/2}x_0, N^{1/2}x_0^2)\right)\end{aligned} \tag{101}$$

yields

$$(m, n)M + \xi = (N^{-1/2}(x_0 + m), N^{1/2}(2mx_0 + n + x_0^2)). \tag{102}$$

Using this result in the definition (97) then confirms the desired (98).

We now follow the same steps as in the previous Section 3 to derive the limiting distribution for $S_N(\ell)$ from equidistribution on $\Gamma\backslash G$. We first state the limit theorem.

THEOREM 4.1. *For any $L > 0$,*

$$\lim_{N\to\infty} \operatorname{meas}\{x_0 \in \mathbb{T} : S_N(\ell) = k\} = E(k, L), \tag{103}$$

where

$$E(k, L) = \frac{1}{\mu(\Gamma\backslash G)} \mu(\bar{g} \in \Gamma\backslash G : F(g) = k), \tag{104}$$

with F as defined in (97).

An explicit formula for $E(0, L)$ and the corresponding gap distribution $P(s)$ (recall Theorem 2.2) is worked out in (Elkies and McMullen, 2004).

The relevant equidistribution theorem needed to prove Theorem 4.1 is the following. Note that $\Gamma \cap \{n_1(x)\}_{x\in\mathbb{R}} = \{n_1(x)\}_{x\in\mathbb{Z}}$ and hence

$$\Gamma\{n_1(x)\}_{x\in\mathbb{T}}\Phi^t \tag{105}$$

represents a family (parametrized by t) of closed orbits embedded in $\Gamma\backslash G$.

THEOREM 4.2. *For any bounded piecewise continuous $f\colon \Gamma\backslash G \to \mathbb{R}$*

$$\lim_{t\to\infty} \int_{\mathbb{T}} f(n_1(x)\Phi^t)\, dx = \frac{1}{\mu(\Gamma\backslash G)} \int_{\Gamma\backslash G} f\, d\mu. \tag{106}$$

Since $n_1(x)$ generates a unipotent flow, Ratner's theorem can be employed. We will explain the general strategy of proof for statements of this type in Section 5.

4.2. SOME HEURISTICS IN THE CASE $\alpha = \sqrt{2}$

We return to generic α, such as $\alpha = \sqrt{2}$, and rewrite $\widetilde{S}_N(\ell)$ as

$$\widetilde{S}_N(\ell) = \sum_{(m,n)\in\mathbb{Z}^2} \chi_{(0,1]}\left(\frac{x_0 + m}{M}\right) \chi_{(-L,L]}\left(\frac{M[\alpha^{-1}(x_0 + m)^2 + n]}{M^{-1}(x_0 + m)}\right) \tag{107}$$

where $M = \sqrt{N\alpha}$. For $x_0 \in [0, 1]$ we can ignore terms of the form x_0/M,

$$\widetilde{S}_N(\ell) \approx \sum_{(m,n)\in\mathbb{Z}^2} \chi_{(0,1]}\left(\frac{m}{M}\right) \chi_{(-L,L]}\left(\frac{M[\alpha^{-1}(x_0 + m)^2 + n]}{M^{-1} m}\right). \tag{108}$$

Now note that for most values of m, we have $m/M \asymp 1$, and it is natural to assume that, for random x_0, the probability of finding k elements of the set

$$\{\alpha^{-1}(x_0+m)^2 : m = 1, \ldots, M\} + \mathbb{Z} \tag{109}$$

in an interval of size $1/M$ around the origin is given by the Poisson distribution (we must assume here that α is badly approximable by rationals, e.g. $\alpha = \sqrt{2}$ would be a good choice). Hence we may assert that the limiting distribution of $S_N(\ell)$ is the same as that of the random variable

$$X = \sum_{(m,n)\in\mathbb{Z}^2} \chi_{(0,1]}\left(\frac{m}{M}\right)\chi_{(-L,L]}\left(\frac{M(\eta_m+n)}{M^{-1}m}\right) \tag{110}$$

where η_m are independent uniformly distributed random variables on $[-\frac{1}{2}, \frac{1}{2})$. With this choice of interval the only contribution comes from the $n = 0$ term (assume $M \gg L$), so

$$X = \sum_{m=1}^{M} X_m \tag{111}$$

where

$$X_m = \chi_{(-L,L]}\left(\frac{M^2\eta_m}{m}\right) \tag{112}$$

is a sequence of independent random variables with kth moment

$$\mathbb{E}X_m^k = \int_{-1/2}^{1/2} \chi_{(-L,L]}\left(\frac{M^2\eta_m}{m}\right) d\eta_m = \frac{2Lm}{M^2}, \tag{113}$$

and hence

$$\mathbb{E}(\mathrm{e}^{\mathrm{i}tX_m} - 1) = \frac{2Lm}{M^2}(\mathrm{e}^{\mathrm{i}t} - 1). \tag{114}$$

The characteristic function of the random variable X is therefore

$$\begin{aligned}
\mathbb{E}\mathrm{e}^{\mathrm{i}tX} &= \prod_{m=1}^{M}\left[1 + \frac{2Lm}{M^2}(\mathrm{e}^{\mathrm{i}t} - 1)\right] \\
&= \exp\left\{\sum_{m=1}^{M} \log\left[1 + \frac{2Lm}{M^2}(\mathrm{e}^{\mathrm{i}t} - 1)\right]\right\} \\
&= \exp\left\{\sum_{m=1}^{M}\left[\frac{2Lm}{M^2}(\mathrm{e}^{\mathrm{i}t} - 1) + O\left(\frac{m^2}{M^4}\right)\right]\right\} \\
&= \exp\left[L(\mathrm{e}^{\mathrm{i}t} - 1) + O\left(\frac{1}{M}\right)\right].
\end{aligned} \tag{115}$$

The expression $\mathrm{e}^{L(\mathrm{e}^{\mathrm{i}t}-1)}$ is the characteristic function of the Poisson law

$$E(k, L) = \frac{L^k}{k!}\mathrm{e}^{-L}. \tag{116}$$

Hence this should be our prediction for the limiting distribution of $S_N(\ell)$, which in turn implies that we expect the exponential distribution for gaps in $\sqrt{m\alpha}$ mod 1. This is in good agreement with our Maple experiment, Figure 2.

5. Ratner's Theorem

An excellent introduction to Ratner's theory is Dave W. Morris' recent textbook (Morris, 2005). Let G be a Lie group (e.g. $\mathrm{SL}(2,\mathbb{R}) \times \mathbb{R}^2$) and Γ be a discrete subgroup (e.g. $\mathrm{SL}(2,\mathbb{Z}) \times \mathbb{Z}^2$). It is at this point not necessary to assume that Γ is a lattice in G, i.e., that $\Gamma\backslash G$ has finite volume with respect to Haar measure μ on G. Ratner's measure classification theorem gives a complete geometric description of all measures that are invariant and ergodic under the a unipotent one parameter subgroup U (or, more generally, invariant and ergodic under a subgroup generated by unipotent subgroups) acting on $\Gamma\backslash G$ by right multiplication. Examples of unipotent subgroups that appeared in the previous sections are $\{n_-(\alpha, 0)\}_{\alpha\in\mathbb{R}}$, $\{n_-(0, y)\}_{\alpha\in\mathbb{R}}$ and $\{n_1(x)\}_{x\in\mathbb{R}}$.

THEOREM 5.1 (Ratner's theorem). *Let ν be an ergodic, U-invariant probability measure on $\Gamma\backslash G$. Then there is a closed, connected subgroup $H \subset G$, and a point $\bar{g} \in \Gamma\backslash G$ such that*

1. *ν is H-invariant,*
2. *ν is supported on the orbit $\bar{g}H$.*

REMARK 5.2. Let $g \in G$ be a representative of the coset $\bar{g} = \Gamma g$, and define the subgroup $\Gamma_H = (g^{-1}\Gamma g) \cap H$. Then the orbit $\bar{g}H$ may be identified with the homogeneous space $\Gamma_H\backslash H$ and ν with the Haar measure on H. Furthermore one can deduce (since ν is a probability measure) that Γ_H is a lattice in H, i.e., $\nu(\bar{g}H) < \infty$, and that the orbit $\bar{g}H$ is closed in $\Gamma\backslash G$.

In simple words, measures ν invariant and ergodic under unipotent subgroups are supported on nice embedded closed subvarieties, of which there can be only countably many (modulo translations of course). We will now discuss two corollaries of Ratner's theorem that are relevant to the equidistribution theorems discussed earlier.

5.1. LIMIT DISTRIBUTIONS OF TRANSLATES

The following is special case of Shah's extremely useful theorem, (Shah, 1996, Theorem 1.4).

```
>  alpha:=sqrt(2); N:=6001;

                              α := √2
                              N := 6001
>  L:=sort([seq(evalf[12](frac(sqrt(n*alpha))), n=1..N)]):
>  alist:=seq(evalf[12](N*(L[i+1]-L[i])),i=1..N-1):
>  data:=stats[transform,tallyinto['outliers']]([alist],[seq((i-1)*
>  i*0.2,i=0..35)]):
>  outliers;

   [7.0547245915, 7.0674227075, 7.1105849000, 7.1693268887, 7.2093775627,
   7.3219323187, 7.3381866273, 7.4195061783, 7.5000233956, 7.6451419780,
   7.7497418084, 7.9388213164, 8.0221013941, 8.1512135092, 8.4582030656]
>  data1:=stats[transform,scaleweight[1/nops([alist])]](data):
>  g1:=stats[statplots,histogram](data1):
>  g2:=plot(exp(-s), s=0..6):
>  plots[display](g1,g2);
```

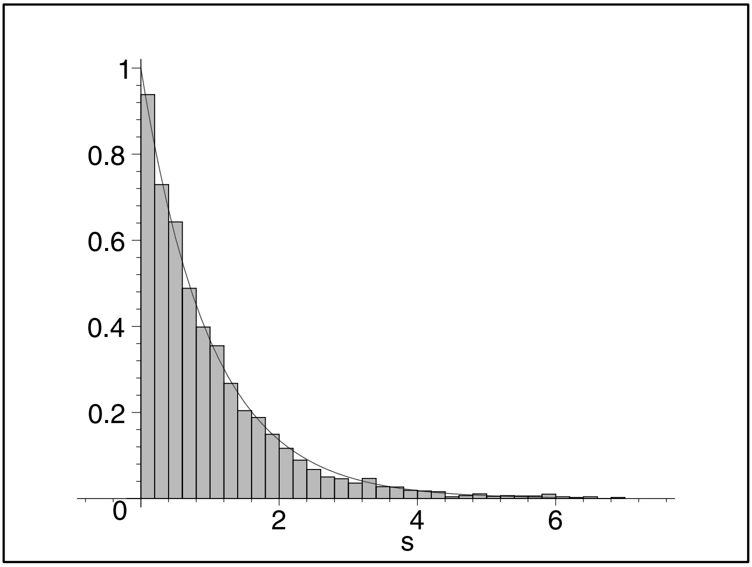


Figure 2. Maple worksheet for calculating the gap distribution of the fractional parts of $\sqrt{m\sqrt{2}}$, $m = 1, \ldots, 6001$.

THEOREM 5.3. *Suppose G contains a Lie subgroup H isomorphic to* $\mathrm{SL}(2,\mathbb{R})$ (*we denote the corresponding embedding by φ*: $\mathrm{SL}(2,\mathbb{R}) \to G$*), such that the set* $\Gamma\backslash\Gamma H$ *is dense in* $\Gamma\backslash G$. *Then, for any bounded, piecewise continuous $f\colon \Gamma\backslash G \to \mathbb{R}$ and any piecewise continuous $h\colon \mathbb{R} \to \mathbb{R}$ with compact support*

$$\lim_{t\to\infty} \int_{\mathbb{R}} f\left(\varphi\left(\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \mathrm{e}^{-t/2} & 0 \\ 0 & \mathrm{e}^{t/2} \end{pmatrix}\right)\right) h(x)\, dx$$

$$= \frac{1}{\mu(\Gamma\backslash G)} \int_{\Gamma\backslash G} f\, d\mu \int_{\mathbb{R}} h(x)\, dx \qquad (117)$$

where μ is the Haar measure of G.

The general strategy of proof for statements of the above type is as follows.

1. Normalize h such that it defines a probability density.
2. Show that the sequences of probability measures ν_t defined by

$$\nu_t(f) = \int_{\mathbb{R}} f\left(\varphi\left(\begin{matrix} 1 & x \\ 0 & 1 \end{matrix}\right.\left(\begin{matrix} \mathrm{e}^{-t/2} & 0 \\ 0 & \mathrm{e}^{t/2} \end{matrix}\right)\right)\right) h(x)\, dx \qquad (118)$$

   is tight. Then, by the Helly–Prokhorov theorem, it is relatively compact, i.e., every sequence of ν_t contains a convergent subsequence with weak limit ν, say.
3. Show that ν is invariant under a unipotent subgroup U; in the present case,

$$U = \left\{\varphi\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\right\}_{x\in\mathbb{R}}. \qquad (119)$$

4. Use a density argument to rule out measures concentrated on subvarieties (exploit the assumption that $\Gamma\backslash\Gamma H$ is dense in $\Gamma\backslash G$).

As an application of Shah's theorem we give a proof of the statement in Remark 3.6, in fact a slightly more general version allowing for non-constant h. Recall that here $G = \mathrm{SL}(2,\mathbb{R}) \times \mathbb{R}^2$ and $\Gamma = \mathrm{SL}(2,\mathbb{Z}) \times \mathbb{Z}^2$.

COROLLARY 5.4. *Let $y \notin \mathbb{Q}$. For any bounded piecewise continuous $f\colon \Gamma\backslash G \to \mathbb{R}$ and piecewise continuous $h\colon \mathbb{R} \to \mathbb{R}$ with compact support*

$$\lim_{t\to\infty} \int_{\mathbb{R}} f(n_{-}(\alpha, y)\Phi^t) h(\alpha)\, d\alpha = \frac{1}{\mu(\Gamma\backslash G)} \int_{\Gamma\backslash G} f d\mu \int_{\mathbb{R}} h(\alpha)\, d\alpha. \qquad (120)$$

Proof. We define the embedding φ: $\mathrm{SL}(2,\mathbb{R}) \to G$ by

$$M \mapsto (1,(0,y))(M,0)(1,(0,y))^{-1}. \tag{121}$$

We need to show that

$$(\gamma,n)(1,(0,y))(M,0)(1,(0,y))^{-1} \tag{122}$$

is dense in G as γ, n, M vary over $\mathrm{SL}(2,\mathbb{Z})$, $\mathbb{Z}^2$, $\mathrm{SL}(2,\mathbb{R})$, respectively. It is obviously sufficient to show this for

$$(\gamma,n)(1,(0,y))(M,0) = (\gamma M,(n_1,(y+n_2))M), \tag{123}$$

and thus for $(M,(n_1,(y+n_2))\gamma^{-1}M)$. It is however easy to see, using the irrationality of y, that $(n_1,(y+n_2))\gamma^{-1}$ is dense in $\mathbb{R}^2$ (exercise). The completes the proof of the density.

Shah's theorem says now that

$$\lim_{t\to\infty}\int_{\mathbb{R}} \tilde{f}(n_-(\alpha,y)\Phi^t n_-(0,y)^{-1})h(\alpha)\,d\alpha = \frac{1}{\mu(\Gamma\backslash G)}\int_{\Gamma\backslash G}\tilde{f}d\mu\int_{\mathbb{R}}h(\alpha)\,d\alpha. \tag{124}$$

for all bounded, piecewise continuous $\tilde{f}$. Choosing the test function

$$\tilde{f}(g) = f(gn_-(0,y)) \tag{125}$$

which is left-Γ-invariant and bounded, piecewise continuous, if f is (as assumed). This yields (120).

5.2. EQUIDISTRIBUTION, UNBOUNDED TEST FUNCTIONS AND DIOPHANTINE CONDITIONS

In some applications of Ratner's theorem, e.g., in questions of value distribution of quadratic forms (Eskin et al., 1998; Eskin et al., 2005; Marklof, 2003; Marklof, 2002), the test functions f in the equidistribution theorems are no longer bounded. Under such circumstances the convergence of the integral can only be assured by assuming certain diophantine conditions. Without going into the intricate details for general $\Gamma\backslash G$, we will illustrate this phenomenon in the distribution of $m\alpha$ on $\mathbb{T}$, which indeed may be viewed as a unipotent orbit on the homogeneous space $\mathbb{Z}\backslash\mathbb{R}$. As mentioned earlier, it is well known that for $\alpha \notin \mathbb{Q}$ the sequence is uniformly distributed mod 1. That is, for any bounded continuous function f: $\mathbb{T} \to \mathbb{R}$ we have

$$\lim_{N\to\infty}\frac{1}{N}\sum_{m=1}^{N} f(m\alpha) = \int_{\mathbb{T}} f(x)\,dx. \tag{126}$$

Let us now formulate the analogous statement for test functions with a possible singularity at $x = 0$.

It is convenient to identify $\mathbb{T}$ with $[-\frac{1}{2}, \frac{1}{2})$. For any $\beta \geq 0$ we define the class $K_\beta(\mathbb{T})$ of functions continuous on $\mathbb{T} - \{0\}$, with the property that there is a constant $C > 0$ such that

$$|f(x)| \leq C|x|^{-\beta}, \quad \text{for all } x \in [-\tfrac{1}{2}, \tfrac{1}{2}). \tag{127}$$

We say $\alpha \in \mathbb{R}$ is *diophantine of type* κ if there exists a constant $c > 0$ such that

$$\left|\alpha - \frac{p}{q}\right| > \frac{c}{q^\kappa}$$

for all $p, q \in \mathbb{Z}$, $q > 0$. The smallest possible value of κ is $\kappa = 2$ (achieved for quadratic surds, e.g., $\alpha = \sqrt{2}$), and it is well known that for any $\kappa > 2$ there is a set of full Lebesgue measure of α that have type κ.

THEOREM 5.5. *Let α be diophantine of type κ. Then, for any $f \in K_\beta(\mathbb{T})$ with* $0 \leq \beta < 1/(\kappa - 1)$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{m=1}^{N} f(m\alpha) = \int_{\mathbb{T}} f(x)\, dx. \tag{128}$$

Proof. We split $f = f_+ - f_-$ into positive and negative part, such that $f_\pm \geq 0$. Then $f_\pm \in K_\beta(\mathbb{T})$ and we may prove (128) for both $f_\pm$ separately. In the following we will thus assume that $f \geq 0$.

For any $\varepsilon > 0$ let us define

$$f_\varepsilon(x) = \begin{cases} f(x) & \text{if } |x| > \varepsilon \\ \min\{f(x), f(\varepsilon)\} & \text{if } |x| \leq \varepsilon \end{cases} \tag{129}$$

and $g_\varepsilon = f - f_\varepsilon$. Note that $f_\varepsilon \leq f$. By assumption,

$$g_\varepsilon(x) \begin{cases} \leq C|x|^{-\beta} & \text{if } |x| \leq \varepsilon \\ = 0 & \text{if } |x| \geq \varepsilon. \end{cases} \tag{130}$$

The function f_ε is bounded continuous, and hence by uniform distribution

$$\lim_{N\to\infty} \frac{1}{N} \sum_{m=1}^{N} f_\varepsilon(m\alpha) = \int_{\mathbb{T}} f_\varepsilon(x)\, dx = \int_{\mathbb{T}} f(x)\, dx - O(\varepsilon^{1-\beta}). \tag{131}$$

Since $f_\varepsilon \leq f$, this implies the lower bound

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{m=1}^{N} f(m\alpha) \geq \int_{\mathbb{T}} f(x)\, dx - O(\varepsilon^{1-\beta}). \tag{132}$$

As to the upper bound,

$$\frac{1}{N}\sum_{m=1}^{N} g_\varepsilon(m\alpha) \le \frac{C}{N}\sum_{m=1}^{N}\frac{\chi_{(0,\varepsilon]}(\|m\alpha\|)}{\|m\alpha\|^\beta} \tag{133}$$

where $\|\cdot\|$ denotes the distance to the nearest integer. Using the dyadic decomposition of the unit interval, we find

$$\begin{aligned}\frac{1}{N}\sum_{m=1}^{N}\frac{\chi_{(0,\varepsilon]}(\|m\alpha\|)}{\|m\alpha\|^\beta} &= \frac{1}{N}\sum_{j=0}^{\infty}\sum_{m=1}^{N}\frac{\chi_{(\varepsilon 2^{-(j+1)},\varepsilon 2^{-j}]}(\|m\alpha\|)}{\|m\alpha\|^\beta}\\ &< \frac{1}{N\varepsilon^\beta}\sum_{j=0}^{\infty}2^{\beta(j+1)}\sum_{m=1}^{N}\chi_{(\varepsilon 2^{-(j+1)},\varepsilon 2^{-j}]}(\|m\alpha\|)\\ &\le \frac{2B}{\varepsilon^\beta}\sum_{j=0}^{\infty}2^{\beta(j+1)}(\varepsilon 2^{-(j+1)})^{1/\kappa-1} \quad \text{(for some } B>0)\\ &= O(\varepsilon^{(1/\kappa-1)-\beta}). \end{aligned}\tag{134}$$

The inequality before the last follows from Lemma 5.6 below. We conclude that

$$\frac{1}{N}\sum_{m=1}^{N} g_\varepsilon(m\alpha) = O(\varepsilon^{(1/\kappa-1)-\beta}) \tag{135}$$

Therefore

$$\begin{aligned}\limsup_{N\to\infty}\frac{1}{N}\sum_{m=1}^{N} f(m\alpha) &= \limsup_{N\to\infty}\frac{1}{N}\sum_{m=1}^{N}[f_\varepsilon(m\alpha)+g_\varepsilon(m\alpha)]\\ &\le \int_{\mathbb{T}} f(x)\,dx + O(\varepsilon^{1-\beta}) + O(\varepsilon^{(1/\kappa-1)-\beta}), \end{aligned}\tag{136}$$

in view of (131) and (135).

Since $\varepsilon > 0$ can be arbitrarily small, the limsup and liminf must coincide.

The following lemma is used in the preceding proof.

LEMMA 5.6. *Let α be diophantine of type κ. Then there is a constant $B > 0$ such that, for any interval $[x_0, x_0 + \ell]$,*

$$\#\{m = 1,\dots,N : m\alpha \in [x_0, x_0+\ell]+\mathbb{Z}\} \le \begin{cases} 0 & \textit{if } N^{\kappa-1}\ell < c\\ BN\ell^{1/(\kappa-1)} & \textit{otherwise.}\end{cases} \tag{137}$$

Proof. Define $T = 1/\ell$. Let us divide the counting into blocks of the form

$$\#\{m_0 < m \le m_0 + T^{1/(\kappa-1)} : m\alpha \in [x_0, x_0 + \ell] + \mathbb{Z}\}, \tag{138}$$

The number of such blocks contributing to (137) is less than $O(NT^{-1/(\kappa-1)} + 1)$.

The gaps between elements of the sequence $m\alpha \bmod 1$, $m_0 < m \le m_0 + T^{1/(\kappa-1)}$, are of the form $n\alpha \bmod 1$, with $|n| < 2T^{1/(\kappa-1)}$. By the diophantine condition, the gaps therefore have seize at least

$$\|n\alpha\| \ge \frac{c}{|n|^{\kappa-1}} > \frac{c}{2^{\kappa-1}T}. \tag{139}$$

An interval of size $\ell = 1/T$ can hence at most contain a bounded number of elements. Hence

$$\#\{m_0 < m \le m_0 + T^{1/(\kappa-1)} : m\alpha \in [x_0, x_0 + \ell] + \mathbb{Z}\} \le B' \tag{140}$$

for some constant $B' > 0$ independent of m_0, x_0, ℓ. Recall that there were at most $NT^{-1/(\kappa-1)}+1$ such blocks, and this yields the upper bound in the second alternative.

The first alternative is easily proven since the minimum gap size for the full sequence $m = 1, \dots, N$ is at least $c/(2N)^{\kappa-1}$.

References

Einsiedler, M., Katok, A., and Lindenstrauss, E. (2006) Invariant measures and the set of exceptions to Littlewoods conjecture, *Ann. of Math. (2)*, to appear.

Elkies, N. D. and McMullen, C. T. (2004) Gaps in $\sqrt{n}$ mod 1 and ergodic theory, *Duke Math. J.* **123**, 95–139.

Eskin, A., Margulis, G., and Mozes, S. (1998) Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture, *Ann. of Math. (2)* **147**, 93–141.

Eskin, A., Margulis, G., and Mozes, S. (2005) Quadratic forms of signature $(2,2)$ and eigenvalue spacings on rectangular 2-tori, *Ann. of Math. (2)* **161**, 679–725.

Kleinbock, D. (1999) Badly approximable systems of affine forms, *J. Number Theory* **79**, 83–102.

Lindenstrauss, E. (2006) Invariant measures and arithmetic quantum unique ergodicity, *Ann. of Math. (2)* **163**, 165–219.

Marklof, J. (2000) The *n*-point correlations between values of a linear form, with an appendix by Z. Rudnick, *Ergodic Theory Dynam. Systems* **20**, 1127–1172.

Marklof, J. (2002) Pair correlation densities of inhomogeneous quadratic forms II, *Duke Math. J.* **115**, 409–434, Correction, *ibid.* **120** (2003) 227-228.

Marklof, J. (2003) Pair correlation densities of inhomogeneous quadratic forms, *Ann. of Math. (2)* **158**, 419–471.

Marklof, J. (2006) Energy level statistics, lattice point problems and almost modular functions, In P. Cartier, B. Julia, P. Moussa, and P. Vanhove (eds.), *Frontiers in Number Theory, Physics and Geometry. Volume 1: On random matrices, zeta functions and dynamical systems*, pp. 163–181, Springer.

Morris, D. W. (2005) *Ratner's theorems on unipotent flows*, Chicago Lectures in Mathematics, Chicago, IL, University of Chicago Press.

Rudnick, Z. and Sarnak, P. (1998) The pair correlation function of fractional parts of polynomials, *Comm. Math. Phys.* **194**, 61–70.

Rudnick, Z. and Zaharescu, A. (2002) The distribution of spacings between fractional parts of lacunary sequences, *Forum Math.* **14**, 691–712.

Shah, N. A. (1996) Limit distributions of expanding translates of certain orbits on homogeneous spaces, *Proc. Indian Acad. Sci., Math. Sci.* **106**, 105–125.

Slater, N. B. (1967) Gaps and steps for the sequence $n\theta \bmod 1$, *Proc. Cambridge Philos. Soc.* **63**, 1115–1123.

Strömbergsson, A. and Venkatesh, A. (2005) Small solutions to linear congruences and Hecke equidistribution, *Acta Arith.* **118**, 41–78.

SPECTRAL THEORY OF AUTOMORPHIC FORMS:
A VERY BRIEF INTRODUCTION

A. Venkatesh
Courant Institute

These are the notes to accompany some lectures delivered at the 2005 NATO ASI summer school in Montréal. They constitute an introduction to the spectral theory of automorphic forms. The viewpoint is slightly nonstandard, in that we present first the "group representation" viewpoint and later descend to the upper-half plane.

The notes reflect quite faithfully the content of the lectures. However I have added references for proofs, which are almost entirely omitted. A major omission is of any form of trace formula: we cover neither the Selberg trace formula, nor the Kuznetsov or Petersson formulae.

1. (Iwaniec, 2002) is the standard introduction to the spectral theory of automorphic forms on the upper half-plane. Includes both Selberg and Kuznetsov formulae.
2. (Borel, 1997) Less oriented to applications in analytic number theory than (Iwaniec, 1987), but perhaps closer to the flavour of these lectures (more representation-theoretic).
3. (Sarnak, 1990) Gives many interesting applications of the spectral theory.
4. (Witte, 2003) An introduction to Ratner's theorems.

1. What Is a Homogeneous Space?

This section is mainly a *teaser trailer* for other people's talks!

1.1. DEFINITION AND EXAMPLES OF HOMOGENEOUS SPACES

The general setting we will be concerned with is the following: G will be a locally compact topological group, $\Gamma \subset G$ a lattice, i.e., a discrete subgroup such that the quotient $X := \Gamma\backslash G$ carries a G-invariant measure of finite volume. The basic question one asks is: if $H \subset G$ is a subgroup, how are H-orbits

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 245–260.

$x_0 H$, for $x_0 \in X$, distributed? One can ask this either in a topological setting (are they dense?) or in a measure setting (are they equidistributed?) Nowadays the strongest results about density are *derived* from proving equidistribution results!

Some examples:

1. The simplest example is $X := \mathbb{Z}\backslash\mathbb{R}$. In this case the dynamics of the $\mathbb{R}$-action on X is the study of circle rotations, which is well-understood.
2. A nilmanifold of depth 2: take

$$\mathbf{U} = \begin{pmatrix} 1 & \star & \star \\ 0 & 1 & \star \\ 0 & 0 & 1 \end{pmatrix}, \quad \Gamma = \mathbf{U}(\mathbb{Z}), \quad G = \mathbf{U}(\mathbb{R}).$$

   The space X is called a *nilmanifold*. Recently a remarkable link has been made between the study of the dynamics of the G-action on nilmanifolds, and finding arithmetic progressions in dense sets of integers (Szemeredi-type theorems) on the other hand: (Host and Kra, 2005; Ziegler, 2006). We do not discuss this further here.

3. A third example, closest to the flavour of what we discuss, is $X := \mathrm{SL}_n(\mathbb{Z})\backslash \mathrm{SL}_n(\mathbb{R})$. In this case the group $G = \mathrm{SL}_n(\mathbb{R})$ is semisimple, and X is not compact, but it nevertheless has finite G-invariant measure. X is identified with the space of rank n unimodular lattices inside $\mathbb{R}^n$.

A common theme of the above examples is that we have started with an algebraic group $\mathbf{G}$ defined over $\mathbb{Q}$, and taken for G the real points $\mathbf{G}(\mathbb{R})$, and for Γ the "integral points" $\mathbf{G}(\mathbb{Z})$. [1] The homogeneous spaces of most relevance to number theory arise in this way: the "number theory" enters through the arithmetic nature of Γ.

1.2. SOME APPLICATIONS TO NUMBER THEORY

Here are some remarkable results of an arithmetic nature that have been derived from equidistribution or density results on homogeneous spaces:

1. A ternary quadratic form represents all sufficiently large squarefree integers that are everywhere locally represented by it. (Duke and Schulze–Pillot (Duke and Schulze-Pillot, 1990), using also some ideas of Iwaniec (Iwaniec, 1987))
2. The Oppenheim conjecture: the values of the quadratic form $x^2 + y^2 - \sqrt{2}z^2$ at $(x, y, z) \in \mathbb{Z}^3$ are dense. (Margulis (Margulis, 1987))

[1] The latter notion is not defined if $\mathbf{G}$ is not defined over $\mathbb{Z}$; one instead embeds $\mathbf{G}$ into SL_M over $\mathbb{Q}$, and takes the intersection $\mathbf{G}(\mathbb{Q}) \cap \mathrm{SL}_M(\mathbb{Z})$.

3. Mazur's conjecture on ranks of elliptic curves over $\mathbb{Z}_p$-extensions: Vatsal (Vatsal, 2002).

Let's briefly explain the connection to homogeneous spaces for the second application. $G = \mathrm{SL}(n, \mathbb{R}), \Gamma = \mathrm{SL}(n, \mathbb{Z})$. It may be verified that Γ is a lattice in G. Moreover, the quotient X is identified with the space of unimodular lattices $L \subset \mathbb{R}^n$ via the map $\Gamma g \mapsto \mathbb{Z}^n \cdot g$. Let Q be a quadratic form on $\mathbb{R}^n$. The set of values taken by Q on $\mathbb{Z}^n$ and $\mathbb{Z}^n.h$ is identical, if h belongs to the isometry group $H := \mathrm{SO}(Q)$. Therefore, if one can prove that the orbit $\mathbb{Z}^n.H$ is dense in the space of unimodular lattices, one verifies the Oppenheim conjecture for Q. Nowadays this would be deduced from a suitable equidistribution statement on X, namely, Ratner's theorem (Ratner, 1991); it was originally proved by Margulis by a direct argument.

1.3. THESE LECTURES: HARMONIC ANALYSIS

Ratner's theorem is proved by ergodic methods. In these lectures I will discuss the harmonic-analysis approach to equidistribution questions on such spaces. Indeed, in the case of $X = \mathbb{Z}\backslash\mathbb{R}$ one has a particularly convenient basis for $L^2(X)$, namely the characters $e_n \colon x \mapsto e^{2\pi i n x}$. They are *convenient* in the following sense: they behave in a very simple way under the $\mathbb{R}$-action.

Our main goal is to construct and discuss the analogous bases for more general homogeneous spaces, i.e., construct a basis for $L^2(\Gamma\backslash G)$ such that the basis functions behave very simply with respect to the G-action. This is possible, at least in part: one does not obtain a basis, but at least a decomposition into much smaller subspaces, which in many ways substitutes for a basis. These subspaces are now indexed by the irreducible representations of G. The study of this will lead us into some basic nonabelian harmonic analysis.

There is a second (perhaps much more important) reason to be interested in the bases we construct: according to Langlands' global reciprocity conjectures, they should parameterize arithmetic objects (e.g., Galois representations). We do not touch on this latter interpretation here.

2. Spectral Theory: Compact Case

2.1. GENERALITIES ON SPECTRAL DECOMPOSITION

In this section I will discuss the case of Γ a cocompact lattice in $G := \mathrm{PSL}(2, \mathbb{R})$. Then G acts by linear fractional transformations on $\mathbb{H}$, the upper half plane. We endow $\mathbb{H}$ with the standard Riemannian metric $(dx^2 + dy^2)/y^2$, with respect to which G acts by isometries. Set K to be the stabilizer in G of

$i \in \mathbb{H}$. Then K is the image in $\mathrm{PSL}(2,\mathbb{R})$ of the subgroup

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \subset \mathrm{SL}(2,\mathbb{R}). \tag{1}$$

Put $Y := \Gamma\backslash\mathbb{H}$. Then $Y = X/K$. Moreover one can identify X to the unit tangent bundle of Y via $\Gamma g \mapsto g(i,\uparrow)$. Here $\uparrow$ is the unit tangent vector at i that corresponds to the direction of the curve parameterized by $t \mapsto i(1+t), t \in \mathbb{R}$ at $t = 0$.

We are trying to generalize the "Fourier decomposition" $L^2(\mathbb{Z}\backslash\mathbb{R}) = \oplus\mathbb{C}e_n$. The really critical property of the $e_n s$ is that they are *eigenfunctions of translation*. Namely, if one translates e_n by t (i.e., considers the function $x \mapsto e_n(x+t)$) one obtains simply $e^{2\pi int}e_n(x)$. One might really hope for the same for $L^2(\Gamma\backslash G)$, i.e., is there a basis f_n of this so that $f_n(xg) = f_n(x)\chi_n(g)$ for all $x \in X, g \in G$ and some suitable functions χ_n? Unfortunately not: the χ_n would have to be characters, and:

EXERCISE 2.1. Prove that the group G has no nontrivial characters, i.e., there do not exist nontrivial homomorphisms from G into the circle group.

To obtain a suitable generalization, we must replace characters of G by higher-dimensional representations.

A unitary representation of G is a continuous isometric action of G on a Hilbert space H, that is to say, a homomorphism ρ of G into the isometries of H with the property that the map $G \times H \to H$ given by $(g,v) \mapsto \rho(g)v$ is continuous. Such a representation is irreducible if any closed G-invariant subspace of H is either trivial, or coincides with H.

Given representations $\pi_i = (H_i,\rho_i)$, one can form in an evident way the direct sum $\oplus_i\rho_i$. There is an evident notion of isomorphism for unitary representations.

EXERCISE 2.2. Show that the representation of G on $L^2(X)$ by right translations, i.e., $\rho(g)f(x) = f(xg)$, defines a unitary representation of G which is not irreducible.

EXERCISE 2.3. Let P be the Lie group of invertible affine transformations $\{x \mapsto ax+b\}$ of the real line, and let P act on $L^2(\mathbb{R})$ via $f \mapsto |a|^{-1/2}f(ax+b)$. Show that this is an irreducible unitary representation.

Let P^+ be the subgroup of P corresponding to elements with $a > 0$. Show that $L^2(\mathbb{R})$ is *not* irreducible as a P^+-representation, but splits as the sum of two irreducible P^+-representations.

Hint for both parts: consider the Fourier transforms. (To avoid topological technicalities, you might like to prove the first assertion first with $\mathbb{R}$ replaced by $\mathbb{Z}/p\mathbb{Z}$).

Let $\widehat{G}$ be the set of isomorphism classes of irreducible unitary representations of the group G.

THEOREM 2.4 (Spectral theorem, compact case). *The representation of G on $L^2(X)$ is isomorphic to a discrete direct sum of irreducible representations, i.e., there are closed subspaces $V_i \subset L^2(X)$ such that*

$$L^2(X) = \widehat{\oplus} V_i$$

where $\widehat{\oplus}$ denotes a Hilbert space direct sum, and each V_i is stable under G and defines an irreducible unitary representation of G.

Moreover, this is discrete *in the sense that given any bounded subset $\Omega \subset \widehat{G}$, there are only finitely many V_i whose isomorphism class belongs to Ω.*

Proof. See (Borel, 1997, Theorem 16.2), which is a proof in the general case of $\Gamma\backslash G$ possibly noncompact.

To make use of this (and to explain the word bounded) we need to understand $\widehat{G}$. Actually, this is not really critically important: one can work out a good (and elegant) theory without having a complete classification like that given below. The point is that it is rare that one needs to know something very explicit about V_i; by and large we only need to know certain properties of the representation (e.g., the asymptotic behavior of matrix coefficients $\langle \rho(g)v_1, v_2 \rangle$ as g varies) and these properties can often be verified without explicit knowledge of V_i.

But for concreteness here we just give the classification, which may seem a little ugly at first sight.

2.2. THE CLASSIFICATION OF IRREDUCIBLE REPRESENTATIONS OF $\mathrm{PSL}_2(\mathbb{R})$

A complete treatment of the representation theory of $\mathrm{SL}_2(\mathbb{R})$ may be found in (Knapp, 2001) (see especially Chapter II) and (Howe and Tan, 1992). It is also contained in Borel's book, (Borel, 1997, Section 15). Note: in this section we present the representation theory of $G = \mathrm{PSL}_2(\mathbb{R})$, i.e., only those representations which are trivial on the center.

A slogan: representations of Lie groups over finite and local fields are to be found in functions on their flag varieties (interpreted liberally, including sections of line bundles over flag varieties). For better slogans see (Vogan, 1998).

In the case of G, the flag variety is the space of lines in $\mathbb{R}^2$ and we shall construct representations in a suitable line bundle over this space.

Let $\nu \in \mathbb{C}$ and let $U(\nu)$ be the space of smooth functions on $\mathbb{R}^2 - \{0\}$ that satisfy $f(\lambda \mathbf{x}) = |\lambda|^{-1-\nu} f(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^2, \lambda \in \mathbb{R}^*$. Then $g \cdot f(x) = f(xg)$ defines an action of G on $U(\nu)$.

If Σ is any nice smooth curve in $\mathbb{R}^2 - \{0\}$ which winds once around the origin, we consider the measure $d\mu_\Sigma$ on Σ that corresponds to area swept out. Then $f \mapsto \int_\Sigma d\mu_\Sigma$ is a G-invariant functional on $U(1)$ which is, in fact, independent of the choice of Σ. (Why?) In particular, the rule $\langle f, g \rangle = \int_\Sigma f \bar{g}\, d\mu_\Sigma$ defines a G-invariant inner product on $U(\nu)$ whenever $\nu \in i\mathbb{R}$. This may be expressed in either of the more concrete forms:

$$\langle f, g \rangle = \frac{1}{2} \int_{S^1} f(\theta) \overline{g(\theta)}\, d\theta = \int_{t \in \mathbb{R}} f(1,t) \overline{g(1,t)}\, dt,$$

where S^1 is the unit circle in $\mathbb{R}^2$, endowed with the angle measure.

Completing $U(\nu)$ w.r.t. this inner product (for $\nu \in i\mathbb{R}$) gives a unitary representation $V(\nu)$ of G.

EXERCISE 2.5. Prove that $V(\nu)$ is isomorphic to the representation of G on $L^2(\mathbb{R})$ defined by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot f(x) = f\left(\frac{b + dx}{a + cx}\right) |a + cx|^{-1-\nu}.$$

Show that $V(\nu)$ is an irreducible unitary representation. (Hint: the restriction of V to upper triangular matrices is closely related to the P^+-representation defined earlier). Note, in particular, that the action of the subgroup $\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ is equivalent to the action of $\mathbb{R}$ on $L^2(\mathbb{R})$ by right translation.

This does not exhaust $\widehat{G}$. In fact, one has the following classes of unitary representations:

1. $V(\nu)$, for $\nu \in i\mathbb{R}$.
2. The trivial representation of G.
3. The *complementary series*: The representations $U(\nu)$ for $\nu \in \mathbb{R}, 0 < \nu < 1$ admit a (non-obvious) G-invariant inner product, and completing leads to the so-called "complementary series" $V(\nu)$.
4. The *discrete series*: For ν an odd integer, each homogeneous polynomial on $\mathbb{R}^2$ of degree $\nu - 1$ belongs to $U(-\nu)$ and so induces a linear functional on $U(\nu)$, via the pairing

$$U(-\nu) \times U(\nu) \xrightarrow{(f,g) \mapsto f.g} U(1) \xrightarrow{f \mapsto \int_\Sigma f d\mu_\Sigma} \mathbb{C}.$$

> The common kernel of these functionals admits a (non-obvious) G-invariant inner product, and completing leads to a representation $V(\nu)$ which decomposes as a sum $V^+(\nu) \oplus V^-(\nu)$.

For simplicity, in these lectures I will only really discuss $V(\nu)$ for $\nu \in i\mathbb{R}$, but for a complete theory one really needs to treat all of $\widehat{G}$.

THEOREM 2.6. *The representations* $\{V(\nu)\}_{\nu \in i\mathbb{R}}$ (*principal series*), $\{V(\nu)\}_{\nu \in (0,1)}$ (*complementary series*), $V^{\pm}(\nu)$ *for* ν *an odd integer* (*discrete series*), *and the trivial representation, are pair-wise nonisomorphic and exhaust* $\widehat{G}$.

2.3. THE SPECTRAL THEOREM ON THE UPPER-HALF PLANE

What is the utility of all the above? We are now capable of reducing questions about the G-action on $L^2(X)$ (which might, *a priori*, depend in some complicated fashion on the lattice Γ) to questions about certain explicit representations that do not know about Γ. A good application of this type of approach will be the proof of mixing of the horocycle and geodesic flows given in Theorem 3.1.

LEMMA 2.7. *Let* π *be an irreducible representation of* G. *The space of* K-*invariant vectors in* π *is one-dimensional if* π *is trivial, principal series, or complementary series, and trivial otherwise.*

Proof. The K-invariant vectors in $U(\nu)$ correspond to rotation-invariant functions. In particular, there is a unique K-invariant vector up to scaling, which corresponds in polar coordinates to the function $r^{-1-\nu}$. It is not too hard to pass to the completion and check that $V(\nu)$ has a unique K-invariant vector up to scaling for $\nu \in (0, 1) \cup i\mathbb{R}$. On the other hand, one can verify that none of the discrete series$V^{\pm}(\nu)$ has a K-invariant vector.

Since $Y = X/K$, we can and will identify K-invariant functions on X with functions on Y. Taking K-invariants in the spectral theorem now yields the spectral theorem in its "upper-half-plane" incarnation.

THEOREM 2.8 (Spectral theorem, compact version, classical form). *Then*

$$L^2(Y) = L^2(X)^K = \oplus_{V_i \text{ not discrete series}} \mathbb{C}\varphi_i$$

where φ_i *is a non-zero* K-*invariant vector in* V_i (*unique up to scaling*). *Moreover, if* V_i *is isomorphic to* $V(\nu_i)$, *then* φ_i *is an eigenfunction of the Laplacian operator* $\Delta := -y^2(\partial_{xx} + \partial_{yy})$ *with eigenvalue* $\frac{1}{4}(1 - \nu^2)$; *and the eigenvalues of the* φ_i *approach* ∞ *as* $i \to \infty$.

Moreover, for $\nu \in i\mathbb{R} \cup (0,1)$*, the* $(1-\nu^2)/4$*-eigenspace of* Δ *is naturally isomorphic to* $\mathrm{Hom}_G(V(\nu), L^2(X))$.

We will discuss how some of the details of this theorem are established in a moment; see also (Iwaniec, 2002, Theorem 4.7), which gives a statement (without the representation-theoretic language) for general—not necessarily cocompact—Γ. For now, note that we have "proved" the (non-obvious) fact that $L^2(Y)$ has an orthonormal basis consisting of eigenfunctions of Δ. Moreover, the eigenspaces exactly parameterize the principal series and complementary series part of $L^2(X)$. To get a corresponding "interpretation" for $\mathrm{Hom}(V^{\pm}(\nu), L^2(X))$ when ν is an odd integer, one considers Laplacian eigenfunctions on suitable line bundles over Y. In this case, the condition of being an eigenfunction is essentially equivalent to being holomorphic, and we are led to the usual theory of holomorphic modular forms.

The main point of Theorem 2.8 is that each φ_i is a Laplacian eigenfunction.

One way to see this is to pass from the action of G to an action of its Lie algebra $\mathfrak{g}$, and thereby to the universal enveloping algebra $U(\mathfrak{g})$. These do not act on $L^2(X)$ but they do act on $C^\infty(X)$. Glossing over this point, each V_i is an *irreducible* representation of G, and by an appropriate version of Schur's lemma the center $\mathfrak{Z}$ of $U(\mathfrak{g})$ acts by a scalar on each V_i. Now this center is generated as a $\mathbb{C}$-algebra by a single element, the Casimir, and explicating everything shows that the fact the Casimir element acts on V_i by a scalar implies that φ_i is a Laplacian eigenfunction.

We sketch another proof that does not require to pass to the Lie algebra and is perhaps more geometrically illuminating. Fix $r \geq 0$. Let Av_r be the operator on functions on the upper half-plane defined by "averaging around circles of radius r", i.e., $\mathrm{Av}_r f(z_0)$, for a function f on $\mathbb{H}$, is the average value of f on an r-circle around $z_0 \in \mathbb{H}$. (Here *average* is taken w.r.t. the length measure on this circle). We claim that, if φ_i is a K-invariant vector in V_i, then $\mathrm{Av}_r\varphi_i$ is a scalar multiple of φ_i. To see this, we note that the function $\mathrm{Av}_r\varphi_i$ corresponds to the function

$$x \mapsto \int_{k\in K} \varphi_i\left(xk\begin{pmatrix} e^{r/2} & 0 \\ 0 & e^{-r/2}\end{pmatrix}\right) dk$$

on X; this function belongs to V_i and is K-invariant, so must be a scalar multiple of φ_i by Lemma 2.7.

Thus φ_i have the following beautiful properties: *they are eigenfunctions of averaging along circles*. (Compare: a harmonic function is one whose average along any circle is zero).

In any case, one can express the Laplacian operator as a limit of the operators Av_r, indeed:

$$\Delta\varphi_i = 4\lim_{r\to 0}\frac{\mathrm{Av}_r\varphi_i - \varphi_i}{r^2}$$

from where it follows that φ_i is also a Δ-eigenfunction (after taking care of suitable issues involving taking the limit).

These two theorems—Theorems 2.4 and Theorem 2.8—give natural bases for $L^2(X)$ and $L^2(Y)$. Actually, we don't have a natural basis for $L^2(X)$, just a canonical decomposition of it into infinitely many subspaces, each of which are infinite dimensional; but this suggests, at the very least, one might try to test the Weyl criterion subspace by subspace.

3. Dynamics

We recall that a one-parameter group $T(t)$ of a probability space (B,μ) is *mixing* if, for any $f, g \in L^\infty(\mu)$, one has $\langle T(t)f, g\rangle_{L^2(\mu)} \to \int f \int g$ as $|t| \to \infty$. Mixing implies ergodicity (i.e., the property that any $T(\mathbb{R})$-invariant set is either zero measure, or has a complement of zero measure). To see this implication, let S be a $T(\mathbb{R})$-invariant set and set $f = g = 1_S$, the characteristic function of S.

THEOREM 3.1. *Let $a(t)$ be the transformation $x \mapsto x\begin{pmatrix} e^t & 0 \\ 0 & e^{-t}\end{pmatrix}$ of X, and $u(t)$ the transformation $x \mapsto x\begin{pmatrix} 1 & t \\ 0 & 1\end{pmatrix}$. Then $a(t)$ and $u(t)$ are mixing (and so ergodic).*

Proof. We sketch part of the proof for $u(t)$, leaving the other case to the reader. By Theorem 2.4, the unitary G-representation $L^2(X)$ decomposes as a direct sum $\oplus_i V_i$. It is easy to see that exactly one of the V_is must consist of the constant functions; say this is $i = 1$. It then suffices to check that, for $f, g \in V_j$ with $j \geq 2$, one has $\langle u(t)f, g\rangle_{L^2(X)} \to 0$ as $t \to \infty$.

We verify this for those j such that V_j is isomorphic to a principal series representation $V(\nu)$. We have seen in Exercise 2.5 that the action of $u(t)$ on $V(\nu)$ is equivalent to translation on $L^2(\mathbb{R})$. The result then follows from the following elementary fact: for $h_1, h_2 \in L^2(\mathbb{R})$, we have

$$\int_{\mathbb{R}} h_1(x)h_2(x+t)\,dx \to 0$$

as $t \to \infty$.

To give a complete proof along these lines, one needs to do a corresponding computation for the complementary series and discrete series. This is indeed doable, but it is also possible to prove this result more directly *without* using the classification of $\widehat{G}$: see, e.g., (Zimmer, 1984, Chapter 2).

In geometric terms, this implies that the geodesic flow and horocycle flow on Y are both mixing. This is very strong: the simplicity of the proof is misleading. To put in perspective the property of mixing, it is worth doing the following exercise:

EXERCISE 3.2. Show that a rotation of the circle is *not* mixing.

From the above Theorem one easily deduces Ratner's theorem in the case of $G = \mathrm{SL}(2, \mathbb{R})$ (but this is again misleading: the general proof of Ratner's theorem is vastly more involved).

4. Spectral Theory: Noncompact Case

The theory of Section 2 does not apply to $\mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{SL}_2(\mathbb{R})$, which is noncompact. It turns out that one cannot have such a simple decomposition for the G-action in this case: one has *continuously occurring families* of irreducible G-representations in $L^2(X)$. I want to avoid the technicalities of setting up this theory, so I will indicate roughly what the results are for $L^2(X)$ but give more details at the level of $L^2(Y)$ instead. The theory of this section applies to a general lattice $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$, but we will discuss only the case of $\mathrm{PSL}_2(\mathbb{Z})$. In this case, the space X (resp. Y) is identified with the space of lattices $L \subset \mathbb{R}^2$ of covolume 1 (resp. such lattices up to rotations).

4.1. CUSPIDAL FUNCTIONS AND EISENSTEIN SERIES ON X

Let N be the upper triangular unipotent subgroup in G, and let $\Gamma_N := N \cap \Gamma$. We say that $f \in L^2(X)$ is cuspidal if $\int_{n\in\Gamma_N\backslash N} f(ng) = 0$ for almost all $g \in G$. The space of cuspidal functions forms a G-invariant closed subspace $L^2_{\mathrm{cusp}}(X) \subset L^2$. Let L^2_{eis} be the orthogonal complement of L^2_{cusp}.

The main results are:

(i) $L^2_{\mathrm{cusp}}(X)$ decomposes as a discrete, direct sum of irreducible G-representations—i.e., it behaves just like $L^2(X)$ in the compact case.

(ii) One can explicitly decompose L^2_{eis} as a G-representation as the direct sum of the trivial representation, and a "direct integral" $\int_{\nu\in i\mathbb{R}} V(\nu)d\nu$ of principal series representations.

The decomposition of L^2_{eis} is much more explicit, in the sense that one can explicitly write a basis in a certain sense. Here we just remark on one fact to give the flavour: For $f \in C^\infty(\mathbb{R}^2)$, construct a function $E_f : X \to \mathbb{R}$, thought of as a function on lattices, via $E_f(L) = \sum_{v \in L} f(v)$. Then $E_f \in L^2_{\mathrm{eis}}$ and such functions span L^2_{eis}. We note that the map $f \mapsto E_f$ is not an L^2-isometry (at least, not for any normal L^2-structure on $C^\infty(\mathbb{R}^2)$) and this is really the main subtlety.

The *general version* of this picture is as follows:

4.2. CUSPIDAL FUNCTIONS AND EISENSTEIN SERIES ON Y

Let us explain this more concretely at the level of $L^2(Y)$. First, given a function $f \in L^2(Y)$, we define the *constant term*

$$f_N(y) := \int_{x \in \mathbb{R}/\mathbb{Z}} f(x + iy)\, dx$$

and we let $L^2_{\mathrm{cusp}}(Y)$ be the subspace of $f \in L^2(Y)$ such that $f_N = 0$.

Let $s \in \mathbb{C}$ have real part larger than 1. If L is any lattice in $\mathbb{R}^2$, the series $\sum_{v \in L} \langle v, v \rangle^{-s}$ is convergent, and it also depends only on L up to rotation. So it defines a function E_s on Y, which can be described also as

$$E_s(z) = \frac{1}{2} \sum_{(c,d) \in \mathbb{Z}^2, (c,d)=1} \frac{y^s}{|cz + d|^{2s}}.$$

THEOREM 4.1. *The function $(y, s) \to E_s(y)$ extends to a meromorphic[2] function on $Y \times \mathbb{C}$ and satisfies $\Delta E_s = s(1 - s)E_s$. The constant term*

$$E_{s,N}(y) = y^s + c(s) y^{1-s},$$

where $c(s) = \xi(2 - 2s)\xi(2s)^{-1}$. Here $\xi(s) = \pi^{-s/2}\Gamma(s)\zeta(s)$. Moreover, E_s satisfies the functional equation $c(s)E_s = E_{1-s}$.

The function E_s is holomorphic when $\Re(s) = 1/2$ and the map

$$g \mapsto \int_{t=0}^{\infty} g(t) E_{1/2+it}\, dt$$

extends to a scalar multiple of a unitary isometry of $L^2(\mathbb{R}_{\geq 0})$ with a subspace $L^2_E \subset L^2(Y)$.

Finally, $L^2(Y) = L^2_{\mathrm{cusp}}(Y) \oplus \langle 1 \rangle \oplus L^2_E$ is an orthogonal decomposition, and $L^2_{\mathrm{cusp}}(Y)$ has a basis consisting of Δ-eigenfunctions whose eigenvalues approach ∞.

[2] It needs to be understood here what "meromorphic" means: here, it is meromorphic in any reasonable sense. For example, given $(y_0, s_0) \in Y \times \mathbb{C}$ there is $N \geq 0$ so that $(s - s_0)^N E_s(y)$ is continuous near (y_0, s_0) and holomorphic in s for each fixed y near y_0.

See (Iwaniec, 2002, Chapter 7) for this theory in the context of more general subgroups of $\mathrm{PSL}_2(\mathbb{R})$.

5. Hecke Operators

We continue with $X = \mathrm{PSL}(2,\mathbb{Z})\backslash \mathrm{PSL}(2,\mathbb{R})$, $Y = \mathrm{PSL}(2,\mathbb{Z})\backslash\mathbb{H}$. One distinctive feature[3] of X is the existence of so-called Hecke operators. In words, there is a large family of "naturally defined" commuting endomorphisms of $C^\infty(X)$ that commute with the G-action. At the level of Y, there is a large family of commuting endomorphisms of $C^\infty(Y)$ that commute with the Laplacian Δ.

Recall that X is identified with the space of lattices in $\mathbb{R}^2$ of covolume 1. For each $x \in X$ let L_x be the corresponding lattice. The lattice L_x has precisely $d(n)$ sublattices of index n, where $d(n)$ is the sum of the divisors of n; each of them has volume n, so scaling them by $n^{-1/2}$ gives a collection $\{L_1, \ldots, L_{d(n)}\} \subset X$. We call this the n-Hecke orbit of L and denote it by $T_n(x) \subset X$. This induces an endomorphism of $L^2(X)$, namely, $T_n f = \sum_{y \in T_n(x)} f(y)$.

Moreover, one has the following three properties of lattices, easily verified:

1. If $L' \subset L$ has index n, then $L'g \subset Lg$ has index n, for any $g \in \mathrm{GL}_2(\mathbb{R})$.
2. If $L' \subset L$ has index n, then $nL \subset L'$ and has index n.
3. Suppose $(n,m) = 1$. The map from chains $L_2 \subset L_1 \subset L$, where $[L : L_1] = n, [L_1 : L_2] = m$, to sublattices $L_2 \subset L$ of index nm, is a bijection.

These translate to the following properties of the endomorphisms T_n of $L^2(X)$:

1. T_n commutes with G;
2. T_n is self-adjoint;
3. $T_n T_m = T_m T_n = T_{nm}$ for $(n,m) = 1$.

One can refine the third property to see that T_n and T_m commute for *all* n and m.

We can refine the spectral decomposition of $L^2(X)$ taking into account the T_ns. I discuss only the cuspidal part. We have already noted (Sec. 4.1) that one may write $L^2(X)_{\mathrm{cusp}} = \oplus V_i$, where each V_i is an irreducible G-representation. However, one can carry out this decomposition in such a way that each T_n (for $n \in \mathbb{N}$) preserves each V_i; by an appropriate form of "Schur's lemma," it

[3] Here "distinctive" means: relative to a quotient of $\mathrm{PSL}(2,\mathbb{R})$ by a "generic" lattice.

follows that T_n acts on V_i by a scalar $\lambda_{i,n} \in \mathbb{R}$. Similarly, one can simultaneously decompose the action of G and $\{T_n\}$ on the orthogonal complement of L^2_{cusp}—this can be carried out even more explicitly—and together this gives a complete spectral resolution of the action of G and the Hecke operators.

The *Ramanujan conjecture* is the assertion that, for all $\varepsilon > 0$, there exists $c = c(\varepsilon)$ such that $|\lambda_{i,n}| \leq cn^{1/2+\varepsilon}$. The strongest known approximation to this is the result of Kim and Sarnak: $|\lambda_{i,n}| \leq c(\varepsilon)n^{1/2+7/64+\varepsilon}$. Bounds on the $|\lambda_{i,n}|$ are related to an equidistribution problem via Weyl's criterion. Indeed, any bound of the form $|\lambda_{i,n}| \ll n^{1-\delta}$, for any fixed positive δ, imply that $\sum_{x \in T_n(y)} \varphi_i(y)/d(n) \to 0$ as $n \to \infty$. This implies that

$$\frac{\sum_{x \in T_n(y)} f(y)}{d(n)} \overset{n\to\infty}{\longrightarrow} \int f, \quad (f \in L^2_{\text{cusp}}). \tag{2}$$

One can extend the validity of (2) from L^2_{cusp} to all of L^2 by using an explicit analysis of the orthogonal complement of L^2 (cf. Sec. 4.1). Thus, *given the bound* $|\lambda_{i,n}| \ll n^{1-\delta}$, one gets the equidistribution result:

THEOREM 5.1. *For $x \in X$, the collection of points $T_n(x)$ becomes equidistributed* (*with respect to the* $\mathrm{PSL}_2(\mathbb{R})$*-invariant measure on X*) *as $n \to \infty$.*

Even this modest result is not easy. To get a rough idea of its content in concrete terms, note that the T_n-orbit of the identity coset in $\mathrm{PSL}_2(\mathbb{Z})\backslash \mathrm{PSL}_2(\mathbb{R})$ is equal to

$$\mathrm{PSL}_2(\mathbb{Z})\backslash \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{Z},\ ad - bc = n \right\}.$$

To understand the behavior of this orbit is closely related to being able to understand the behavior of integral solutions to $q(a, b, c, d) = n$, where q is the quaternary quadratic form $q(a, b, c, d) = ad - bc$. The analysis of quaternary quadratic forms over $\mathbb{Z}$ was already successfully carried out by Kloosterman, using his refinement of the Hardy–Littlewood method and a nontrivial bound for Kloosterman sums (Kloosterman, 1926); and, indeed, Theorem 5.1 is of exactly the same character, and can be deduced from a nontrivial bound for Kloosterman sums.

(We remark that Theorem 5.1 is also very closely related to the following result, that was discussed in Duke's lectures: the integral solutions to $x^2 + y^2 + z^2 = n^2$ are equidistributed (when projected to the sphere of radius 1) as $n \to \infty$. Again, this result follows from any nontrivial bound on Kloosterman sums.)

We conclude with an amusing application of the Theorem which relates to some lectures of Rudnick during this meeting. Choose a random pair $(a, b) \in$

$(\mathbb{Z}/p\mathbb{Z})^2 - \{0,0\}$. Let $N(a,b)$ be the smallest value of $\sqrt{x^2+y^2}$ when (x,y) range through all nonzero integral solutions to the congruence $(x,y) \equiv t(a,b) \mod p$, for some $t \in \mathbb{Z}$. Then (as $p \to \infty$) the distribution of $N(a,b)/\sqrt{p}$, when (a,b) is chosen at random, coincides with the distribution of $1/\sqrt{y}$, when $x+iy$ is chosen at random, w.r.t. the hyperbolic measure, from the standard fundamental domain $\{z : |z| \geq 1, |\Re(z)| \leq 1/2\}$ for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$.

EXERCISE 5.2. Explain the above using Hecke operators. What happens if one replaces $\sqrt{x^2+y^2}$ by $\max(|x|,|y|)$?

For more in this line see (Strömbergsson and Venkatesh, 2005).

6. Gross Omissions: The Selberg Trace Formula

We have omitted any discussion of the trace formula. From the viewpoint of group representation theory, to decompose a representation of a group, one should compute its character. So what is the "character" of the G-representation $L^2(\Gamma\backslash G)$? The answer to this is given by the "trace formula"; the only slight hitch is the trace does not make sense as a function on G, and needs to be regularized. However, the trace does make sense as a *distribution*, and this is what the trace formula computes.

Baby case: think about the action of a translation operator on $L^2(\mathbb{Z}\backslash\mathbb{R})$, i.e., consider for $\alpha \in \mathbb{R}$ the automorphism R_α that sends a periodic function f to $x \mapsto f(x+\alpha)$. How would you define its trace? It is "diagonal" w.r.t. the orthonormal basis $\{e^{2\pi i n x}\}_{n\in\mathbb{Z}}$ for $L^2(\mathbb{Z}\backslash\mathbb{R})$. Therefore, we are naively led to define its trace as the infinite sum $\sum_n e^{2\pi i n\alpha}$. This is divergent for all α; we need a way to regularize it. To do this, we note that this infinite sum makes sense *as a distribution*. That is to say: $\sum_n e^{2\pi i n\alpha}$ makes no sense as a function, but it defines a perfectly valid distribution, namely, the following functional on $C^\infty(\mathbb{R})$:

$$g(\alpha) \mapsto \sum_n \int_{\alpha\in\mathbb{R}} g(\alpha)e^{2\pi i n\alpha}. \tag{3}$$

We leave it to the reader to verify that the right hand side is convergent and defines a continuous functional on C^∞, w.r.t. the usual topology. But this is also a very well-known distribution: the Poisson summation formula shows that it equals $\sum_{n\in\mathbb{Z}} g(n)$. Thus, the *distributional* trace of $L^2(\mathbb{Z}\backslash\mathbb{R})$ is the sum of Dirac measures, one at each integer.

The *trace formula*, then, computes the "distributional trace" of G acting on $L^2(\Gamma\backslash G)$. In the case that Γ is cocompact in G, it says *precisely* what one would expect by generalizing the character formula for an induced representation, familiar from finite groups. (Recall that if $H_1 \subset H_2$ are finite groups, the H_2 representation on $L^2(H_1\backslash H_2)$ is, in more usual terminology, the induction to H_2 of the trivial representation on H_1). In particular, the distributional trace of this representation is supported on conjugacy classes of G that contain an element of Γ.

When one unravels everything in this case of Γ cocompact, the trace formula gives (among other things) an identity of the following type: if h is a nice test function, and $\lambda_1, \dots, \lambda_r, \dots$ the eigenvalues of Δ on $L^2_{\mathrm{cusp}}(\Gamma\backslash\mathbb{H})$, then

$$\sum_i h(\lambda_i) = \sum_{\gamma\in\Gamma^\sharp} \hat{h}(\mathrm{Trace}(\gamma)) \tag{4}$$

where:

1. $\hat{h}$ is a transformed version of h;
2. $\Gamma^\sharp$ is the set of conjugacy classes in Γ (which are in correspondence with closed geodesics on $Y = \Gamma\backslash\mathbb{H}$.)

From (4) one may deduce many results about the distribution of the λ_i, such as (Weyl's law) the asymptotic for $\#\{\lambda_i \leq X\}$ as $X \to \infty$. It should be noted that Weyl's law is far more elementary, however, and (4) gives much more precise information.

In the case when Γ is not cocompact, matters are *much* more complicated; you can see the problem with noncompact spaces by trying to replace $L^2(\mathbb{Z}\backslash\mathbb{R})$ by $L^2(\mathbb{R})$ in the above discussion and see what happens. In the case when Γ is not cocompact, then, one computes the trace not of the G-action on $L^2(\Gamma\backslash G)$, but of the G-action on the subspace $L^2_{\mathrm{cusp}}(\Gamma\backslash G)$. However, the trace formula yields the following fundamental result: the space of *even* cusp forms in $L^2_{\mathrm{cusp}}(\mathrm{PSL}_2(\mathbb{Z})\backslash\mathbb{H})$ is nonzero, and indeed infinite dimensional. Here *even* means: preserved by the symmetry $x + iy \mapsto -x + iy$.

Some references: (Iwaniec, 2002, Chapter 10) (classical viewpoint, no representation theory, emphasis on analytic applications) (Gelbart and Jacquet, 1979) (adelic and representation-theoretic viewpoint for GL_2, emphasis on functorial applications), (Gelbart, 1996) (adelic viewpoint, introduction to trace formulae on bigger groups).

Acknowledgements

I would like to thank the CRM, and especially the organizers Andrew Granville and Zeév Rudnick, for the 2005 Nato ASI summer school. It was a very pleasant experience, certainly for the lecturers, and, hopefully, for some of the students also! I would also like to acknowledge the support of the Clay Mathematics Institute, of the Institute for Advanced Study (NSF Grants DMS-0111298), and of NSF Grant DMS-0245606.

References

Borel, A. (1997) *Automorphic forms on* $\mathrm{SL}_2(\mathbb{R})$, Vol. 130 of *Cambridge Tracts in Math.*, Cambridge Univ. Press.

Borel, A. and Jacquet, H. (1979) Automorphic forms and automorphic representations, In *Automorphic forms, representations and L-functions*, Vol. 33 of *Proc. Sympos. Pure Math.*, Corvallis, OR, 1977, pp. 189–202, Providence, RI, Amer. Math. Soc.

Duke, W. (1988) Hyperbolic distribution problems and half-integral weight Maass forms, *Invent. Math.* **92**, 73–90.

Duke, W., Friedlander, J., and Iwaniec, H. (1993a) Bounds for automorphic L-functions, *Invent. Math.* **112**, 1–8.

Duke, W., Friedlander, J., and Iwaniec, H. (1993b) Bounds for automorphic L-functions, *Invent. Math.* **112**, 1–8.

Duke, W. and Schulze-Pillot, R. (1990) Representation of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids, *Invent. Math.* **99**, 49–57.

Gelbart, S. (1996) *Lectures on the Arthur–Selberg trace formula*, Vol. 9 of *Univ. Lecture Ser.*, Providence, RI, Amer. Math. Soc.

Gelbart, S. and Jacquet, H. (1979) Forms of GL(2) from the analytic point of view, *Proc. Sympos. Pure Math.* **33**, 213–251.

Host, B. and Kra, B. (2005) Nonconventional ergodic averages and nilmanifolds, *Ann. of Math.* **161**, 398–488.

Howe, R. and Tan, E.-C. (1992) *Nonabelian harmonic analysis. Applications of* $\mathrm{SL}(2,\mathbb{R})$, Universitext, New York, Springer.

Iwaniec, H. (1987) Fourier coefficients of modular forms of half-integral weight, *Invent. Math.* **87**, 385–401.

Iwaniec, H. (2002) *Spectral methods of automorphic forms*, Vol. 53 of *Grad. Stud. Math.*, Providence, RI, Amer. Math. Soc.

Kloosterman, H. D. (1926) On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$, *Acta Math.* **49**, 407–464.

Knapp, A. (2001) *Representation theory of semisimple groups. An overview based on examples*, Princeton, NJ, Princeton Univ. Press.

Margulis, G. (1987) Formes quadratriques indéfinies et flots unipotents sur les espaces homogènes, *C. R. Acad. Sci. Paris Sér. I Math.* **304**, 249–253.

Ratner, M. (1991) Raghunathan's topological conjecture and distributions of unipotent flows, *Duke Math. J.* **63**, 249–253.

Sarnak, P. (1990) *Some applications of modular forms*, Vol. 99 of *Cambridge Tracts in Math.*, Cambridge Univ. Press.

Strömbergsson, A. and Venkatesh, A. (2005) Small solutions to linear congruences and Hecke equidistribution, *Acta Arith.* **118**, 41–78.

Vatsal, V. (2002) Uniform distribution of Heegner points, *Invent. Math.* **148**, 1–46.

Vogan, D. (1998) The method of coadjoint orbits for real reductive groups, In *Representation theory of Lie groups*, Park City, UT, 1998, pp. 179–258, Providence, RI, Amer. Math. Soc.

Witte, D. (2003) Ratner's theorems on unipotent flows, arXiv: math.DS/0310402.

Ziegler, T. (2006) Universal characteristic factors and Furstenberg averages, *J. Amer. Math.Soc.*, to appear; arXiv:math.DS/0403212.

Zimmer, R. (1984) *Ergodic theory and semisimple groups*, Vol. 81 of *Monographs Math.*, Birkhäuser.

SOME EXAMPLES HOW TO USE MEASURE CLASSIFICATION IN NUMBER THEORY

Elon Lindenstrauss
Princeton University

Abstract. We give examples of how classifying invariant probability measures for specific algebraic actions can be used to prove density and equidistribution results in number theory.

1. Introduction

1.1.

Ergodic theory has proven itself to be a powerful method to tackle difficult number theoretical problems, particularly problems which involve equidistribution.

A typical application involves three parts: (a) translating the number theoretical problem into a problem about specific algebraically defined actions; (b) classifying invariant measures for the action; (c) deducing the desired equidistribution statement from this measure classification.

1.2.

All the actions we will consider are of the following form: the space on which the action takes place is a quotient space $X = \Gamma\backslash G$ where G is a linear algebraic group, and $\Gamma < G$ a lattice[1]. Any subgroup H of the group of affine transformations[2] on G mapping Γ-cosets to Γ-cosets acts on X. In particular, any subgroup $H < G$ acts on X by right translations $h.x = xh^{-1}$.

This is a fairly broad class of actions. Typically, for specific number theoretic applications one needs to consider a specific action. For example, in §3 we give a proof due to Furstenberg of the equidistribution of $n^2\alpha$ mod 1 by studying the $\mathbb{Z}$ action generated by the affine map $(x, y) \mapsto (x + \alpha, y + 2x + \alpha)$ on the space $X = \mathbb{R}^2/\mathbb{Z}^2$. Margulis proved the long-standing Oppenheim

[1] i.e., a discrete subgroup of finite covolume.

[2] i.e., the group of maps $G \to G$ generated by right translations and automorphisms.

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 261–303.

conjecture by studying the action of SO(2, 1), i.e., the group of linear transformations preserving a fixed indefinite quadratic form (say $x_1^2 + x_2^2 - x_3^2$) in three variables, on $X = \mathrm{SL}(3, \mathbb{Z})\backslash \mathrm{SL}(3, \mathbb{R})$, the space of covolume one lattices in $\mathbb{R}^3$. In §6 we present results from (Einsiedler et al., 2004) towards Littlewood's conjecture regarding simultaneous Diophantine approximations by studying the action of the group of 3×3 diagonal matrices of determinant one on the same space $X = \mathrm{SL}(3, \mathbb{Z})\backslash \mathrm{SL}(3, \mathbb{R})$.

1.3.

Weyl's (nonergodic) original proof the equidistribution of $n^2\alpha$ mod 1 is not very complicated and, unlike most ergodic theoretic methods, gives quantitative results regarding equidistribution rates (see, e.g., (Granville and Rudnick, 2006) in these proceedings); but the elegance of Furstenberg's proof is quite striking. Furthermore, it is a good illustration of the general scheme discussed in §1.1 and serves as a simple model for the other, more complicated, results we discuss and quote later.

1.4.

A very general measure classification theorem which lies at the heart of numerous deep number theoretical applications is Ratner's measure classification theorem (§4.6). We discuss this theorem, and particularly how it can be applied to prove equidistribution in §4. Returning to the general scheme presented in §1.1, the first step of translating a number theoretic question to one related to dynamics seems to be more of an art than a science. The second step is provided by Ratner's measure classification theorem, which is a deep and complicated theorem, but which can sometimes be used as a black box. The reader could certainly profit from learning some of the ideas involved but this is beyond the scope of this paper; besides, the recent book (Morris, 2005) seems to cover these ideas quite well. Therefore we focus on the third part of the general scheme, namely how to use this measure classification effectively; we discuss in particular some results and techniques of (Dani and Margulis, 1993) that seem to be particularly useful in this respect. Even with regard to this part, we only attempt to illustrate clearly the issues that need to be addressed; the reader who really wants to master this important technique should study in detail one of the papers quoted where such an application is carried out.

Ratner's theorem and how to apply it are also considered from a somewhat different perspective in (Markloff, 2006) in this volume.

1.5.

Ratner's theorem does not cover all algebraic actions which arise naturally from number theoretic problems. A good example is the action of the full diagonal group on $\mathrm{SL}(n,\mathbb{Z})\backslash \mathrm{SL}(n,\mathbb{R})$.

A good understanding of entropy theory is absolutely essential to applying what results we have regarding invariant measures for these actions to number theory. Therefore we devote considerable space in §5 to present some of the fundamentals regarding entropy[3].

1.6.

Next we present in detail in §6 one application due to Einsiedler, Katok and the author (Einsiedler et al., 2004) of a partial measure classification result to estimating the set of possible exceptions to Littlewood's conjecture. As in §4 we focus on how a measure classification result (Einsiedler et al., 2004, Theorem 1.3) is applied and not on the measure classification result itself. Our treatment is quite close to that of (Einsiedler et al., 2004) though some of the results are presented in a slightly more explicit form.

1.7.

Finally, in §7 we explain how measure classification is related to the behavior of Laplacian eigenfunctions on arithmetic quotient spaces—specifically the arithmetic quantum unique ergodicity question. At first sight the measure classification problem one is led to does not seem to be a promising one as there are too many invariant measures, but hidden symmetries and restrictions save the day.

1.8.

This paper was written with a fairly narrow aim: to aid graduate students who are interested in understanding the interplay between ergodic theory and number theory and are willing to spend some effort doing so. Of course, other people may find this paper helpful or at least entertaining.

This paper is certainly not a survey, in the traditional sense of the word. Some topics are given detailed even technical treatment, while some are discussed only superficially. It is certainly not meant to be comprehensive and the choice of topics is fairly subjective and arbitrary. While I have made some

[3] The serious reader would do well to study entropy beyond what we provide here, e.g. from (Rudolph, 1990).

effort to give correct attributions, doubtless some inaccuracies remain—the reader interested in a detailed historical account should look elsewhere.

It is the author's intention to continue updating this tutorial, and eventually to publish an expanded and more detailed version elsewhere. As it is, it already contains quite a bit of material and (supplemented with pertinent references) can probably be used as a basis for a one semester graduate course on homogeneous dynamics and applications.

1.9.

A word about notations: the paragraphs in this paper are numbered, and this numbering is logically identified with the numbering of theorems, definitions, etc.; e.g., "Ratner's measure classification theorem (§4.6)" and "Theorem 4.6" are synonyms. Hopefully this will survive the typesetting.

2. Dynamical Systems: Some Background

2.1.

DEFINITION. Let X be a locally compact space, equipped with an action of a noncompact (but locally compact) group[4] H. An H-invariant probability measure μ on X is said to be *ergodic* if one of the following equivalent conditions holds:

(i) Suppose $A \subset X$ is an H-invariant set, i.e., $h.A = A$ for every $h \in H$. Then $\mu(A) = 0$ or $\mu(A^{\complement}) = 0$.

(ii) Suppose f is a measurable function on X with the property that for every $h \in H$, for μ-a.e. x, $f(h.x) = f(x)$. Then f is constant a.e.

(iii) μ is an extreme point of the convex set of all H-invariant Borel probability measures on X.

2.2.

A stronger condition which implies ergodicity is mixing:

DEFINITION. Let X, H and μ be as in Definition 2.1. The action of H is said to be mixing if for any sequence $h_i \to \infty$ in H[5] and any measurable subsets $A, B \subset X$,

$$\mu(A \cap h_i.B) \to \mu(A)\mu(B) \quad \text{as } i \to \infty.$$

[4] All groups will be assumed to be second countable locally compact, all measures Borel probability measures unless otherwise specified.

[5] i.e. a sequence so that for any compact $C \subset H$ only finitely many of the h_i are in C.

2.3.

A basic fact about H-invariant measures is that any H-invariant measure is an average of ergodic measures, i.e., there is some auxiliary probability space (Ξ, ν) and a (measurable) map attaching to each $\xi \in \Xi$ an H-invariant and ergodic probability measure μ_ξ on X so that

$$\mu = \int_\Xi \mu_\xi \, d\nu(\xi).$$

2.4.

DEFINITION. An action of a group H on a locally compact topological space X is said to be uniquely ergodic if there is only one H-invariant probability measure on X.

2.5.

The simplest example of a uniquely ergodic transformation is the map $T_\alpha \colon x \mapsto x + \alpha$ on the one dimensional torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ where α is irrational. Clearly Lebesgue measure m on $\mathbb{T}$ is T_α-invariant; we need to show it is the only such probability measure.

To prove this, let μ be an arbitrary T_α-invariant probability measure. Since μ is T_α-invariant,

$$\hat{\mu}(n) = \int_\mathbb{T} e(nx) \, d\mu(x) = \int_\mathbb{T} e(n(x + \alpha)) \, d\mu(x) = e(n\alpha)\hat{\mu}(n),$$

where as usual $e(x) = \exp(2\pi i x)$. Since α is irrational, $e(n\alpha) \neq 1$ for all $n \neq 0$, hence $\hat{\mu}(n) = 0$ for all $n \neq 0$ and $\mu = m$.

2.6.

DEFINITION. Let X be a locally compact space, and suppose that $H = \{h_t\} \cong \mathbb{R}$ acts continuously on X. Let μ be a H-invariant measure on X. We say that $x \in X$ is *generic for μ* if for every $f \in C_0(X)$ we have[6]:

[6] Where $C_0(X)$ denotes the space of continuous functions on X which decay at infinity, i.e., so that for any $\varepsilon > 0$ the set $\{x \colon |f(x)| \geq \varepsilon\}$ is compact.

$$\frac{1}{T}\int_0^T f(h_t.x)\,dt \to \int_X f(y)\,d\mu(y) \quad \text{as } T \to \infty.$$

Equidistribution is another closely related notion:

2.7.

DEFINITION. A sequence of probability measures μ_n on a locally compact space X is said to be *equidistributed* with respect to a (usually implicit) measure m if they converge to m in the weak* topology, i.e., if $\int f\,d\mu_n \to \int f\,dm$ for every $f \in C_c(X)$.

A sequence of points $\{x_n\}$ in X is said to be equidistributed if the sequence of probability measures $\mu_N = N^{-1}\sum_{n=1}^{N}\delta_{x_n}$ is equidistributed, i.e., if for every $f \in C_0(X)$

$$\frac{1}{N}\sum_{n=1}^{N} f(x_n) \to \int_X f(y)\,dm(y) \quad \text{as } N \to \infty.$$

Clearly there is a lot of overlap between the two definitions, and in many situations“ equidistributed” and “generic” can be used interchangeably.

2.8.

For an arbitrary $H \cong \mathbb{R}$-invariant measure μ on X, the Birkhoff pointwise ergodic theorem shows that μ-almost every point $x \in X$ is generic with respect to some H-invariant and ergodic probability measure on X. If μ is ergodic, μ-a.e. $x \in X$ is generic for μ.

If X is compact, and if the action of $H \cong \mathbb{R}$ on X is uniquely ergodic with μ being the unique H-invariant measure, then something much stronger is true: *every* $x \in X$ is generic for μ!

Indeed, let μ_T be the probability measures

$$\mu_T = \frac{1}{T}\int_0^T \delta_{h_t.x}\,dt$$

then any weak* limit of the μ_T will be H-invariant. But there is only one H-invariant probability measure on X, namely μ, so $\mu_T \to \mu$, i.e., x is generic for μ.

3. Equidistribution of $n^2\alpha$ mod 1

3.1.

A famous theorem of Weyl states that for any irrational α, the sequence $n^2\alpha$ mod 1 is equidistributed. In this section we give an alternative proof, due to Furstenberg, which proves this theorem by classifying invariant measures on a suitable dynamical system. We follow Furstenberg's treatment in (Furstenberg, 1981, §3.3).

3.2.

The dynamical system we will study is the following: the space will simply be the 2-torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, and the action will be the one generated by the map

$$T\colon (x,y) \mapsto (x+\alpha, y+2x+\alpha). \tag{1}$$

One easily proves using induction that

$$T^n(x,y) = (x+n\alpha, y+2nx+n^2\alpha). \tag{2}$$

We will prove below that $(\mathbb{T}^2, T)$ is uniquely ergodic. By §2.8, it follows that for *every* (x,y) the orbit $\{T^n(x,y)\}_{n=1}^{\infty}$ is equidistributed. In particular, the orbit of the point $(0,0)$ is equidistributed, i.e.

$$\{(n\alpha \text{ mod } 1, n^2\alpha \text{ mod } 1)\colon n \in \mathbb{N}\}$$

is equidistributed. We see that unique egodicity of T implies not only equidistribution of $n^2\alpha$ mod 1 but the stronger fact that $(n\alpha \text{ mod } 1, n^2\alpha \text{ mod } 1)$ is equidistributed in $\mathbb{T}^2$.

The same proof, with minor modifications, can be used to show equidistribution of $p(n)$ mod 1 for any polynomial with an irrational leading coefficient (see Exercise 3.8 below).

3.3.

The proof that (X,T) is uniquely ergodic is harder than for irrational rotations on $\mathbb{T}$ (cf. §2.5). The basic scheme, which is not unusual in such proofs, is that we first prove that Lebesgue measure m on $\mathbb{T}^2$, which is obviously invariant under T, is also ergodic. A separate argument is then used to bootstrap the ergodicity of Lebesgue measure to unique ergodicity.

3.4.

PROPOSITION. *Lebesgue measure m on $\mathbb{T}^2$ is ergodic under T.*

Proof. Let $f \in L^2(m)$ be T-invariant. Expand f to a Fourier series

$$f(x, y) = \sum_{n,m} \hat{f}_{n,m} e(nx + my).$$

By T-invariance,

$$\hat{f}_{n,m} = \hat{f}_{n+2m,m} e((n + m)\alpha). \tag{3}$$

In particular, $\left|\hat{f}_{n,m}\right| = \left|\hat{f}_{n+2m,m}\right|$. By the Riemann–Lebesgue lemma, $\hat{f}_{n,m} \to 0$ as $(n, m) \to \infty$, hence $\hat{f}_{n,m} = 0$ if $m \neq 0$.

For $m = 0$, however, (3) becomes $\hat{f}_{n,0} = e(n\alpha)\hat{f}_{n,0}$, so $\hat{f}_{n,m} = 0$ for all $(n, m) \neq 0$.

It follows that f is constant a.e., and m is ergodic.

This argument cannot be applied directly for T-invariant probability measures, as the Fourier transform of a measure does not satisfy the Riemann–Lebesgue Lemma.

3.5.

The bootstrapping argument which we use to upgrade ergodicity to unique ergodicity is a simple positivity argument.

PROPOSITION. *Let g be a measurable function $\mathbb{T} \to \mathbb{T}$, and $T_g \colon \mathbb{T}^2 \to \mathbb{T}^2$ be the map*

$$T_g(x, y) = (x + \alpha, y + g(x))$$

with α irrational. Then if the Lebesgue measure m is T_g-ergodic, then in fact it is the only T_g-invariant probability measure, i.e., $(\mathbb{T}^2, T_g)$ is uniquely ergodic.

Proof. Suppose $\mu \neq m$ is another T_g-invariant probability measure on $\mathbb{T}^2$. Let R_a denote the map $(x, y) \mapsto (x, y + a)$. Then since T_g and R_a commute, for any $a \in \mathbb{T}$, $(R_a)_*\mu$ is also T_g-invariant. Consider the measure

$$m' = \int_{\mathbb{T}^2} (R_a)_*\mu. \tag{4}$$

Clearly m' is invariant under R_a for every a, and its projection to the first coordinate has to be a probability measure invariant under the rotation $x \mapsto x + \alpha$, hence Lebesgue. It follows that $m' = m$. But by assumption, m is ergodic, and hence is an extreme point in the convex set of T_g invariant probability measures on $\mathbb{T}^2$. Therefore m cannot be presented as a nontrivial linear combination of other T_g invariant probability measures, in contradiction to (4).

3.6.

Proposition 3.4 and Proposition 3.5 together clearly imply

COROLLARY. *The map* $T\colon (x,y) \mapsto (x+\alpha, y+2x+\alpha)$ *on* $\mathbb{T}^2$ *is uniquely ergodic for every irrational* α*.*

As discussed in §3.2, equidistribution of $\{n^2\alpha \bmod 1\}$ is now an easy consequence of this corollary.

3.7.

The proof we have given for the equidistribution of $\{n^2\alpha \bmod 1\}$ is very elegant, but it has one serious drawback compared to Weyl's original method: it does not give rates. The ambitious reader is encouraged to try and figure out how to modify Furstenberg's proof to obtain a more quantitative result regarding the rate of equidistribution. Such a quantification of a qualitative ergodic theoretic argument is often referred to as effectivization, and often can be quite entertaining and worthwhile.

3.8.

EXERCISE. Generalize this argument to give an ergodic theoretic proof for the equidistribution of $p(n)$ mod 1 for any polynomial $p(n)$ with an irrational leading coefficient.

4. Unipotent Flows and Ratner's Theorems

4.1.

A very general and important measure classification theorem has been proved by Ratner, in response to conjectures by Dani and Raghunathan. For simplicity, we restrict our treatments to the case of Lie groups, even though the extension of Ratner's theorems to products of real and p-adic groups (Ratner, 1995; Margulis and Tomanov, 1994) is just as important for number theoretical applications; for a recent and striking example, see (Ellenberg and Venkatesh, 2006)).

4.2.

DEFINITION. An element $g \in \mathrm{GL}(n,\mathbb{R})$ is said to be *unipotent* if all its (real or complex) eigenvalues are equal to one. An element g in a Lie group G is said to be Ad-*unipotent* if $\mathrm{Ad}(g)$ is a unipotent element of $\mathrm{GL}(\mathfrak{g})$, with $\mathfrak{g}$ the Lie algebra of G.

4.3.

DEFINITION. Let X be a topological space, and H a locally compact group acting continuously on X. An orbit $H.x$ is said to be *periodic* if it has a finite H-invariant measure.[7]

EXAMPLE. Suppose $H = \mathbb{Z}$, and that the action of H on X is generated by the map $T: X \to X$. Then x has a periodic H-orbit iff $T^n x = x$ for some $n \in \mathbb{N}$.

4.4.

We remark that in the locally homogeneous context, i.e., $X = \Gamma\backslash G$ and H a subgroup of G acting on X by right translations, every periodic H-orbit is also a closed subset of X (Raghunathan, 1972, Theorem 1.13).

4.5.

Let G be a Lie group, $\Gamma < G$ a discrete subgroup, and $H < G$. One obvious class of H invariant probability measures on $\Gamma\backslash G$ are L-invariant probability measures on single periodic L-orbits for (closed) subgroups $L < G$ containing H. We shall call such measures *homogeneous*; equally common in this context is the adjective *algebraic*.

4.6.

THEOREM (Ratner's measure classification theorem (Ratner, 1991a)). *Let G be a Lie group, $\Gamma < G$ a discrete subgroup, and $H < G$ a closed connected subgroup generated by* Ad-*unipotent one parameter groups. Then any H-invariant and ergodic probability measure on $\Gamma\backslash G$ is homogeneous* (*in the sense of* §4.5).

While the statement of Theorem 4.6 the group Γ is not assumed to be a lattice[8], for most applications this assumption is necessary, as otherwise the assumption that the H-invariant measure under consideration is a *probability* measure is not a natural one.

For the remainder of this section, unless otherwise specified, Γ will be a *lattice* in G.

[7] More formally, there is a nontrivial finite H-invariant measure ν on X so that $\nu(X - H.x) = 0$.

[8] i.e., a discrete subgroup of finite covolume

4.7.

The proof of Theorem 4.6 is beyond the scope of this paper. The ambitious reader is encouraged to study the proof; helpful references are the recent book (Morris, 2005) (particularly Chapter 1), Ratner's treatment of a "baby case" in (Ratner, 1992), and a simplified self contained proof of the special case $H \cong \mathrm{SL}(2,\mathbb{R})$ (but general G and Γ) in (Einsiedler, 2006). A more advanced reference (in addition to Ratner's original papers) is Margulis and Tomanov's proof of this result (Margulis and Tomanov, 1994) which in particular uses entropy theory as a substitute to some of Ratner's arguments. A useful survey paper which covers much of what we discuss in this section is (Kleinbock et al., 2002), particularly (Kleinbock et al., 2002, §3).

4.8.

Consider for simplicity first the case of H itself a unipotent one parameter flow (in particular, as an abstract group, $H \cong \mathbb{R}$). If there are no H-invariant probability measures on $\Gamma\backslash G$ other than the G-invariant measure, then as we have seen in §2.8 it is fairly straightforward to deduce from the measure classification theorem information regarding how each individual orbit is distributed, and in particular classify the possible orbit closures (which in the uniquely ergodic case can be only $\Gamma\backslash G$ itself, i.e., the H-flow is *minimal*).

4.9.

EXERCISE.

(i) Let $G = \mathrm{SL}(2,\mathbb{R})$ and $\Gamma < G$ a *cocompact* lattice. Let H be the group $\left\{\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}\right\}$. Deduce from Ratner's theorem that the action H on $\Gamma\backslash G$ is uniquely ergodic.

(ii) Let $H < G$ be as in (i), but take $\Gamma = \mathrm{SL}(2,\mathbb{Z})$. What are the H-invariant measures in this case?

The proof that H acting on $\Gamma\backslash G$ in (i) above is uniquely ergodic predated Ratner's theorem by about 20 years and is due to Furstenberg (Furstenberg, 1973). The classification in (ii) is due to Dani (Dani, 1978). There are much simpler proofs now, the simplest (and quite quantitative) proof is via the mixing of the geodesics flow.

4.10.

In most cases of interest, however, there is more than one invariant measure in Theorem 4.6, and in this case deducing information about individual orbits from a measure classification theorem is far less obvious (cf. Exercise 4.15 below). Nonetheless measure classification is the main ingredient in the proof of the following two important results of Ratner:

4.11.

THEOREM (Ratner's genericity theorem (Ratner, 1991b)). *Let G be a Lie group, Γ a lattice, and H a unipotent one parameter subgroup of G. Then* every $x \in \Gamma\backslash G$ *is generic for a homogeneous measure supported on a periodic orbit $L.y$ containing x.*

4.12.

THEOREM (Ratner's orbit closure classification theorem (Ratner, 1991b)). *Let G be a Lie group, Γ be a lattice in G, and $H < G$ closed connected subgroup generated by* Ad-*unipotent one parameter groups. Then for any $x \in \Gamma\backslash G$, the orbit closure $\overline{H.x}$ is a single periodic orbit for some group $H \leq L \leq G$.*

4.13.

Ratner's theorems give us very good understanding of the dynamics of groups generated by unipotents on finite volume quotients $\Gamma\backslash G$. Much less is known about the case when Γ is a discrete subgroup with infinite covolume. For example, we do not know how to classify Radon measures[9] invariant under unipotent groups in the infinite covolume case, and we do not understand orbit closures in this case except in very special cases (for example, see (Burger, 1990; Ledrappier and Sarig, 2005; Roblin, 2003)).

The action of groups H which are not generated by unipotents on $\Gamma\backslash G$ (even in the finite covolume case) is also not well understood at present. This topic will be discussed in detail in §6.

[9] i.e., locally finite but possibly infinite measures.

4.14.

Ratner's proof of Theorem 4.12 via measure classification is not the only approach to classifying orbit closures. In particular, for the important special case of $H = \mathrm{SO}(2,1)$[10] Dani and Margulis (Dani and Margulis, 1989) (following earlier work of Margulis) classified all possible orbit closures $\overline{H.x}$ in $\mathrm{SL}(3,\mathbb{Z})\backslash\mathrm{SL}(3,\mathbb{R})$ before Ratner's work.

The action of this group $H = \mathrm{SO}(2,1)$ on $\mathrm{SL}(3,\mathbb{Z})\backslash\mathrm{SL}(3,\mathbb{R})$ is closely connected to the Oppenheim conjecture regarding values of indefinite quadratic forms which was posed in the 1920's, and was only solved in the 1980's by Margulis (see, e.g., (Margulis, 1989)) using a partial classification of orbit closures of this action. An accessible self-contained treatment of this result is (Dani and Margulis, 1990).

It is not clear (at least to me) exactly what is the limit of these topological methods, and whether they can be pushed to give a full proof of Theorem 4.12.

4.15.

EXERCISE. Let $X = \{0,1\}^{\mathbb{Z}}$ and $\sigma\colon X \to X$ be the shift map $(\sigma(x))_i = x_{i+1}$. Let $n_i \uparrow \infty$ be an increasing sequence of integers with $n_i/i \to 0$. Define Y to be the set of those sequences $x \in X$ with the property that for every $i \in \mathbb{N}$, the sequence "01" does not appear more than i times in any stretch of n_i+1-digits, i.e., if $B = \{x \in X\colon x_0 = 0, x_1 = 1\}$,

$$Y = \left\{ x \in X\colon \forall i \in \mathbb{N}, \max_j \sum_{k=j}^{j+n_i-1} 1_B(\sigma^k x) \leq i \right\}.$$

(i) Show that there are precisely two σ-invariant and ergodic probability measures on Y. What are they?

(ii) Prove that there are $y \in Y$ which are not generic for any σ-invariant probability measure on Y.

(iii) Show that there are uncountably many possible orbit closures for σ (i.e., sets of the form $\overline{\{\sigma^n y\colon n \in \mathbb{Z}\}}$ with $y \in Y$).

[10] i.e., the group of determinant one matrices preserving a fixed quadratic form of signature 2,1—e.g. $Q(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_3^2$.

4.16.

Exercise 4.15 shows that one cannot deduce Ratner's strong rigidity statements about individual orbits from her measure classification theorem by purely formal means. This reduction is carried out by Ratner in (Ratner, 1991b). Dani and Margulis (Dani and Margulis, 1993) subsequently gave a somewhat different treatment which gives more uniform and flexible versions of Theorem 4.11 that are often highly useful in number theoretic applications.

In many applications, the level of detail we give is unnecessary. For example, in (Vatsal, 2002), Theorem 4.12 was precisely what was needed (and a fairly simple case of this theorem at that). Markloff in his contribution to this proceeding (Markloff, 2006) deduces interesting results about statistical properties of some number theoretic sequences from the results of (Shah, 1996) (which are proved using the same principles exposed in this section). The results of (Mozes and Shah, 1995) also seem to be very relevant, and in particular almost immediately imply the equidistribution statements proved in Exercise 4.28. However, it is my belief that the serious user of Ratner's theorem should have some understanding of what is involved in the reduction of equidistribution statements to Ratner's measure classification theorem (more so than the inner workings of this measure classification theorem, which are probably more interesting for the ergodic theorist wanting to push this technology further).

4.17.

The main difficulty in passing from a measure classification theorem to a theorem about behavior of individual orbits is that orbits may for some stretch of time behave according to some invariant measure, and then after a relatively short transition period start behaving according to a different invariant measure.

There is an extra difficulty in the locally homogeneous context in that more often than not the space we consider is not compact, bringing in another complication: to pass from measure classification to statements regarding individual orbits one needs to be able to control how much time an orbit spends far away (i.e., outside big compact sets). Both of these difficulties (which are closely related) can be addressed by the following basic estimates.

4.18.

DEFINITION. Let G be a Lie group and $\Gamma < G$ a discrete subgroup. For any subgroup $H < G$ define the *singular set relative to* H, denoted by $\mathcal{S}(H)$, as the union of all periodic orbits in $X = \Gamma\backslash G$ of all closed subgroups $L < G$ containing H.

If H is a one parameter Ad-unipotent group then by Theorem 4.11, $\mathcal{S}(H)$ is precisely the set of all $x \in X$ which fail to be generic for the G-invariant measure on X with respect to the action of H.

It is worthwhile to delve a bit into the structure of this singular set $\mathcal{S}(H)$. Suppose that $L_1.x$ is a periodic orbit with $L_1 > H$ and $x = \pi_\Gamma(g)$. Let $L = {}^gL_1$ where we use the notations ${}^gL = gLg^{-1}$ and $L^g = g^{-1}Lg$. Then since $L_1.x$ has finite volume, $\Gamma_L = L \cap \Gamma$ is a lattice in L. Let

$$X(L, H) = \{h \in G\colon hHh^{-1} \subset L\}.$$

For any $h \in G$, since $L_1.x$ is periodic, the orbit of $y = \pi_\Gamma(h)$ under $L_2 = {}^{h^{-1}}L$ is periodic. If $h \in X({}^gL, H)$, we have that the natural probability measure on this periodic orbit $L_2.y$ is H-invariant. In this way we get a "tube" of periodic orbits $\pi_{\Gamma_L}(X(L, H))$ on $\Gamma_L\backslash G$ which descends to a family of periodic orbits $\pi_\Gamma(X(L, H))$ on X. Of course, for some L and H this family may be empty or consist of a single periodic orbit.

By (Dani and Margulis, 1993, Proposition 2.3),

$$\mathcal{S}(H) = \bigcup_{L\in\mathcal{H}} \pi_\Gamma(X(L, H))$$

where $\mathcal{H}$ is a countable collection of closed connected subgroups of G.[11]

EXERCISE. Work this decomposition out explicitly for $G = \mathrm{SL}(2, \mathbb{R})$, $\Gamma = \mathrm{SL}(2, \mathbb{Z})$, and $H = \left\{\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}\right\}$.

4.19.

A careful understanding of this singular set (cf. (Dani and Margulis, 1990, §3)) is important to control the amount of time a unipotent trajectory can spend near a lower dimensional invariant subspace. For instance, it can be used to show the following:

THEOREM ((Dani and Margulis, 1993, Theorem 1)). *Let H be a closed connected subgroup of G generated by* Ad-*unipotent elements. Let $F \subset X - \mathcal{S}(H)$*

[11] Namely, the collection of all closed connected subgroups $L < G$ satisfying that (a) $\dim L < \dim G$, (b) $L \cap \Gamma$ is a lattice in L, and (c) the image of $\Gamma \cap L$ under the adjoint representation is Zariski dense in the image of L.

be compact. Then for any $\varepsilon > 0$, there is a neighborhood Ω of $\mathcal{S}(H)$ such that for any Ad*-unipotent one parameter subgroup $\{u_t\}$ of G, any $x \in F$, and any $T \geq 0$,*

$$\mathrm{Leb}(\{t \in [0, T] \colon u_t.x \in \Omega\}) \leq \varepsilon T.$$

Note that typically $\mathcal{S}(H)$ is dense so the condition on F above is rather harsh. A more precise result is (Dani and Margulis, 1993, Theorem 7.3) which essentially shows that the only way the trajectory $\{u_t.\pi_\Gamma(g) \colon t \in [0, T]\}$ spends a substantial amount of time near the singular set is if g is so close to some $X(L_0, H)$ so that *for all $t \in [0, T]$* the point gu_{-t} is close to $X(L_0, H)$.

4.20.

In order to control the related question of how much time an arbitrary orbit of the one parameter unipotent subgroup spends in a neighborhood of infinity we have the following, which follows from several papers of Dani and Margulis starting with (Margulis, 1971).

THEOREM ((Dani and Margulis, 1993, Theorem 6.1)). *Let G be a Lie group and $\Gamma < G$ a lattice. Then for any compact $F \subset \Gamma\backslash G$ and any $\varepsilon > 0$ there is a compact $C \subset X$ so that for any* Ad*-unipotent one parameter subgroup $\{u_t\}$ of G, any $x \in F$ and any $T > 0$*

$$\mathrm{Leb}(\{t \in [0, T] \colon u_t.x \notin C\}) \leq \varepsilon T.$$

There are many extensions and variations on this result, some of them quite important. A nice place to read about some of these developments (and indeed also about the basic method) is (Kleinbock and Margulis, 1998).

4.21.

As a basic example of how Theorems 4.19 and 4.20 can be used, we show how with the aid of these theorems, Ratner's theorem about generic points (Theorem 4.11) can be deduced from her measure classification theorem (Theorem 4.6)[12].

Proof. Let $x \in X$, and for any $T > 0$ set μ_T to be the probability measure

$$\mu_T = \frac{1}{T} \int_0^T \delta_{u_t.x} \, dt.$$

Without loss of generality, we may assume that $x \notin \mathcal{S}(\{u_t\})$ for otherwise we may replace X (and G and Γ accordingly) with a (lower dimensional) periodic orbit containing $\{u_t.x\}$.

[12] This is not how Ratner proved Theorem 4.11!

What we want to prove is that for any $f \in C_0(X)$

$$\int_X f\, d\mu_T \overset{?}{\rightarrow} \int_X f\, dm \quad \text{as } T \to \infty \tag{5}$$

where m is the G invariant probability measure on X, i.e., that μ_T converge weak* to m.

By Theorem 4.20 applied to $F = \{x\}$, for any $\varepsilon > 0$ there is a compact set $C \subset X$ so that for all T we have $\mu_T(C) > 1 - \varepsilon$. It follows that there is a sequence of $T_i \uparrow \infty$ for which μ_{T_i} converge in the weak* topology to a *probability* measure μ_∞.

This limiting measure μ_∞ is invariant under u_t. By Theorem 4.6 and the ergodic decomposition, μ_∞ is a linear combination of m and the natural probability measures on periodic orbits of groups L containing H. In particular, $\mu_\infty = \alpha m + (1 - \alpha)\mu'$ with μ' a probability measure on $\mathcal{S}(\{u_t\})$.

Applying Theorem 4.19 to $F = \{x\}$, we get for any $\varepsilon > 0$ an open set $\Omega \supset \mathcal{S}(\{u_t\})$ with $\mu_T(\Omega) < \varepsilon$ for all T. It follows that

$$\mu_\infty(\Omega) \leq \varliminf_{i \to \infty} \mu_{T_i}(\Omega) \leq \varepsilon$$

and so $\alpha \geq 1 - \varepsilon$. Since ε was arbitrary, we see that $\mu_\infty = m$ and Theorem 4.11 follows.

4.22.

EXERCISE. Use Theorem 4.20 to show that (5) holds for any continuous bounded f (not necessarily decaying at infinity). This slightly stronger form of Theorem 4.11 is the one given in (Ratner, 1991b).

4.23.

EXERCISE. Use a similar arguments to prove the following ((Dani and Margulis, 1990, Theorem 2)):

Let $u_t, u_t^{(1)}, u_t^{(2)}, \dots$ be one parameter Ad-*unipotent subgroups of G with $u_t^{(i)} \to u_t$, x_i be a sequence of points in X converging to $x \in X - \mathcal{S}(\{u_t\})$, and $T_i \uparrow \infty$. Then for any continuous bounded f*

$$\frac{1}{T_i} \int_0^{T_i} f(u_t^{(i)}.x_i)\, dt \to \int_X f\, dm.$$

Hint: show first that without loss of generality we can assume $x_i \notin \mathcal{S}(\{u_t\})$.

4.24.

We end this section with an interesting application, presented in the form of an exercise, of Ratner's theorems and the related results of Dani–Margulis to equidistribution of the points of Hecke correspondences. This application was first suggested by Burger and Sarnak (Burger and Sarnak, 1991) and a detailed proof was given by Dani and Margulis in (Dani and Margulis, 1993). Recently Eskin and Oh (Eskin and Oh, 2006a) gave a further generalization of this approach.

All these results are quite general, but we consider only the simplest case of $X = \Gamma\backslash \mathrm{SL}(2, \mathbb{R})$. This case is also discussed in Venkatesh' contribution to these proceedings (Venkatesh, 2006).

4.25.

We begin our discussion by defining the Hecke correspondences for the case of $G = \mathrm{SL}(2, \mathbb{R})$, $\Gamma = \mathrm{SL}(2, \mathbb{Z})$; as usual let $X = \Gamma\backslash G$. We say that an integer matrix $\gamma \in M_2(\mathbb{Z})$ is irreducible if there is no nontrivial integer dividing all its coefficients.

DEFINITION. Let n be an integer ≥ 2. The *n-Hecke correspondence* is a map which assigns to a point $x = \pi_\Gamma(g) \in X$ a finite subset $T_n(x)$ of X (with the number of points in $T_n(x)$ depending only on n) by

$$T_n(x) = \{\pi_\Gamma(n^{-1/2}\gamma g) \colon \gamma \in M_2(\mathbb{Z}) \text{ irreducible with } \det \gamma = n\}.$$

While it is not completely obvious from the formula, $T_n(x)$ is a finite collection of points of X, and its cardinality can be given explicitly and depends only on n.

Using the Hecke correspondence we can define operators (also denoted T_n) on $L^2(X)$ by

$$T_n(f)[x] = c_n \sum_{y \in T_n(x)} f(y);$$

where we take[13] $c_n = |T_n(x)|^{-1}$.

[13] This is not the standard normalization; the standard normalization is $c_n = n^{-1/2}$.

4.26.

For example[14], if $n = p$ is prime, and $x = \pi_\Gamma(g)$ as above, $T_p(x)$ consists of the $p + 1$ points

$$T_p(x) = \left\{\pi_\Gamma\left(p^{-1/2}\begin{pmatrix} p & 0 \\ 0 & 1\end{pmatrix} g\right), \pi_\Gamma\left(p^{-1/2}\begin{pmatrix} p & 0 \\ 1 & 1\end{pmatrix} g\right), \ldots, \right.$$
$$\left. \pi_\Gamma\left(p^{-1/2}\begin{pmatrix} p & 0 \\ p-1 & 1\end{pmatrix} g\right), \pi_\Gamma\left(p^{-1/2}\begin{pmatrix} 1 & 0 \\ 0 & p\end{pmatrix} g\right)\right\};$$

and for any n, the set $T_{n^2}(x)$, which contains roughly n^2 points, contains in particular the $\phi(n)$ points

$$\pi_\Gamma\left(\begin{pmatrix} 1 & k/n \\ 0 & 1\end{pmatrix} g\right), \quad (k, n) = 1.$$

The operators T_n respects the action of $G = \mathrm{SL}(2, \mathbb{R})$ by translations on X, i.e., $T_n(g.x) = g.T_n(x)$. For future reference, we note that this shows that the Hecke operators T_n descend to correspondences on $\Gamma\backslash\mathbb{H} \cong \Gamma\backslash G/K$ with $K = \mathrm{SO}(2, \mathbb{R})$.

4.27.

The following theorem was discussed in Venkatesh' contribution to these proceedings:

THEOREM. *For any $x \in X$, the points of the Hecke correspondences $T_n(x)$ become equidistributed*[15] *as $n \to \infty$.*

There a spectral approach to the theorem is discussed, using the known bounds towards the Ramanujan Conjecture, which gives much sharper results than what one can presently get using ergodic theory. However, it is quite instructive to deduce this equidistribution statement from Ratner's theorems.

4.28.

EXERCISE. Let $G^2 = G \times G$, $\Gamma^2 = \Gamma \times \Gamma$, $X^2 = \Gamma^2\backslash G^2$ and $G_\Delta < G^2$ the subgroup $G_\Delta = \{(g, g) : g \in G\}$. Also let $u(t) = \begin{pmatrix} 1 & t \\ 0 & 1\end{pmatrix}$ and $u_\Delta(t) = (u(t), u(t))$. Let m_X and denote the G invariant measure on X and similarly for X^2.

[14] Which the reader should verify!

[15] To be more precise, the sequence of probability measures $|T_n(x)|^{-1} \sum_{y \in T_n(x)} \delta_y$ become equidistributed in a sense of Definition 2.7.

The purpose of this exercise[16] is to prove the following, which is essentially equivalent[17] to the equidistribution of $T_{n^2}(x_0)$ for every $x_0 \in X$ as $n \to \infty$ along the lines of (Burger and Sarnak, 1991; Dani and Margulis, 1993)

For any $f, g \in L^2(X)$, we have that

$$\int T_{n^2}(f) g \, dm_X \to \int f \, dm_X \int g \, dm_X \quad \text{as } n \to \infty. \tag{6}$$

Let $\alpha \in (0,1) - \mathbb{Q}$ be arbitrary, and let $k(n)$ be a sequence of integers satisfying (a) $(k(n), n) = 1$ and (b) $k(n)/n \to \alpha$.

(i) Let $y_n^2 = \pi_{\Gamma^2}\left(\begin{pmatrix} 1 & k(n)/n \\ 0 & 1 \end{pmatrix}, e\right)$, with e denoting the identity. Show that

$$G_\Delta . y_n^2 = \{(x,y) : x \in X, y \in T_{n^2}(x)\}.$$

Let $y_\infty^2 = \pi_{\Gamma^2}\left(\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}, e\right)$.

(ii) Let μ_n denote the natural measure on the periodic orbit $G_\Delta . y_n^2$. Show, using the well-known ergodicity of the action of $u(t)$ on $\Lambda \backslash G$ (for any lattice Λ), that $u_\Delta(t)$ acts ergodically on μ_n.

(iii) Deduce from Theorem 4.20, the fact that $y_n^2 \to y_\infty^2$, the ergodicity of $u_\Delta(t)$ acting on μ_n and the pointwise ergodic theorem that μ_n converge weak* to a probability measure μ.

(iv) Show that $y_\infty^2 \notin \mathcal{S}(G_\Delta)$, and deduce similarly from Theorem 4.19 that $\mu(\mathcal{S}(G_\Delta)) = 0$.

(v) Use Ratner's measure classification theorem (§4.6) to deduce $\mu = m_{X^2}$.

(vi) Consider $F(x,y) = f(x)g(y)$. Show that

$$\int F(x,y) \, d\mu_n = \int T_{n^2}(f) g \, dm_X.$$

Deduce (6).

[16] This exercise is somewhat advanced and may use more than we assume in the rest of this paper.

[17] Only in this particular instance, because of the equivariance of T_p under G.

4.29.

The reader is encouraged to look at other applications of Ratner's theorem to equidistribution and counting problems, for example (Eskin et al., 1996; Eskin et al., 1998; Elkies and McMullen, 2004; Eskin and Oh, 2006b).

5. Entropy of Dynamical Systems: Some More Background

5.1.

A very basic and important invariant in ergodic theory is entropy. It can be defined for any action of a (not too pathological) unimodular amenable group H preserving a probability measure (Ornstein and Weiss, 1987), but for our purposes we will only need (and only consider) the case $H \cong \mathbb{R}$ or $H \cong \mathbb{Z}$.

Entropy was lurking behind the scenes already in the study of the action of unipotent groups considered in §4[18], but plays a much more prominent role in the study of diagonalizable actions which we will consider in the next section.

5.2.

Let (X, μ) be a probability space. The entropy $H_\mu(\mathcal{P})$ of a finite or countable partition of X is defined to be

$$H_\mu(\mathcal{P}) = -\sum_{P \in \mathcal{P}} \mu(P) \log \mu(P).$$

One basic property of entropy is sub-additivity; the entropy of the refinement $\mathcal{P} \vee \mathcal{Q} = \{P \cap Q \colon P \in \mathcal{P}, Q \in \mathcal{Q}\}$ satisfies

$$H_\mu(\mathcal{P} \vee \mathcal{Q}) \le H_\mu(\mathcal{P}) + H_\mu(\mathcal{Q}). \tag{7}$$

However, this is just a starting point for many more natural identities and properties of entropy, e.g., equality holds in (7) if and only if $\mathcal{P}$ and $\mathcal{Q}$ are independent.

5.3.

The ergodic theoretic entropy $h_\mu(T)$ associated to a measure preserving map $T \colon X[2] \to X$ can be defined using the entropy function H_μ as follows:

[18] In particular, in (Margulis and Tomanov, 1994) Margulis and Tomanov give a substantially shorter proof of Ratner's measure classification theorem using entropy theory.

DEFINITION. Let μ be a probability measure on X and $T\colon X \to X$ a measurable map preserving μ. Let $\mathcal{P}$ be either a finite partition of X or a countable partition with $H_\mu(\mathcal{P}) < \infty$. The entropy of the four-tuple $(X, \mu, T, \mathcal{P})$ is defined to be[19]

$$h_\mu(T, \mathcal{P}) = \lim_{N\to\infty} \frac{1}{N} H_\mu\left(\bigvee_{n=0}^{N-1} T^{-n}\mathcal{P}\right). \tag{8}$$

The ergodic theoretic entropy of (X, μ, T) is defined to be

$$h_\mu(T) = \sup_{\mathcal{P}: H_\mu(\mathcal{P})<\infty} h_\mu(T, \mathcal{P}).$$

The ergodic theoretic entropy was introduced by A. Kolmogorov and Ya. Sinai and is often called the Kolmogorov–Sinai entropy; it is also somewhat confusingly called the metric entropy (even though it has nothing to do with any metric that might be defined on X!).

If μ is a T-invariant but not necessarily ergodic measure, it can be shown that the entropy of μ is the average of the entropy of its ergodic components: i.e., if μ has the ergodic decomposition $\mu = \int \mu_\xi \, d\nu(\xi)$, then

$$h_\mu(T) = \int h_{\mu_\xi}(T) \, d\nu(\xi). \tag{9}$$

5.4.

A partition $\mathcal{P}$ is said to be a *generating partition* for T and μ if the σ-algebra $\bigvee_{n=-\infty}^{\infty} T^{-n}\mathcal{P}$ (i.e., the σ-algebra generated by the sets $\{T^n P\colon n \in \mathbb{Z}, P \in \mathcal{P}\}$) separates points; that is, for μ-almost every x, the atom of x with respect to this σ-algebra is $\{x\}$.[20] The Kolmogorov–Sinai theorem asserts the non-obvious fact that $h_\mu(T) = h_\mu(T, \mathcal{P})$ whenever $\mathcal{P}$ is a generating partition.

5.5.

We also want to define the ergodic theoretic entropy also for flows (i.e., for actions of groups $H \cong \mathbb{R}$). Suppose $H = \{a_t\}$ is a one parameter group acting on X. Then it can be (fairly easily) shown that for $s \neq 0$, $1/|s| h_\mu(x \mapsto a_s.x)$ is independent of s. We define the entropy of μ with respect to $\{a_t\}$, denoted $h_\mu(a_\bullet)$, to be this common value of $1/|s| h_\mu(x \mapsto a_s.x)$.[21]

[19] Note that by the subadditivity of the entropy function H_μ the limit in (8) exists and is equal to $\inf_N (1/N) H_\mu(\bigvee_{n=0}^{N-1} T^{-n}\mathcal{P})$.

[20] Recall that the atom of x with respect to a countably generated σ-algebra $\mathcal{A}$ is the intersection of all $B \in \mathcal{A}$ containing x and is denoted by $[x]_{\mathcal{A}}$.

[21] Note that $h_\mu(a_\bullet)$ depends not only on H as a group but on the particular parametrization a_t.

5.6.

Suppose now that (X, d) is a compact metric space, and that $T : X \to X$ is a homeomorphism (the pair (X, T) is often implicitly identified with the generated $\mathbb{Z}$-action and is called a dynamical system).

DEFINITION. The $\mathbb{Z}$ action on X generated by T is said to be *expansive* if there is some $\delta > 0$ so that for every $x \neq y \in X$ there is some $n \in \mathbb{Z}$ so that $d(T^n x, T^n y) > \delta$.

If X is expansive then any measurable partition $\mathcal{P}$ of X for which the diameter of every element of the partition is $< \delta$ is generating (with respect to any measure μ) in the sense of §5.4.

5.7.

For the applications presented in the next section, an important fact is that for many dynamical systems (X, T) the map $\mu \mapsto h_\mu(T)$ defined on the space of T-invariant probability measures on X is semicontinuous. This phenomenon is easiest to see when (X, T) is expansive.

PROPOSITION. *Suppose (X, T) is expansive, and that μ_i, μ are T-invariant probability measures on X with $\mu_i \to \mu$ in the weak* topology. Then*

$$h_\mu(T) \geq \overline{\lim_{i\to\infty}}\, h_{\mu_i}(T).$$

In less technical terms, for expansive dynamical systems, a "complicated" invariant measure might be approximated by a sequence of "simple" ones, but not vice versa.

Proof. Let $\mathcal{P}$ be a partition of X such that for each $P \in \mathcal{P}$

(i) $\mu(\partial P) = 0$.

(ii) P has diameter $< \delta$ (δ as in the definition of expansiveness).

Since $\mu(\partial P) = 0$ and $\mu_i \to \mu$ weak*, for every $P \in \mathcal{P}$ we have that $\mu_i(P) \to \mu(P)$. Then

$$\begin{aligned} \frac{1}{N} H_\mu \left(\bigvee_{n=0}^{N-1} T^{-n}\mathcal{P} \right) &= \lim_{i\to\infty} \frac{1}{N} H_{\mu_i} \left(\bigvee_{n=0}^{N-1} T^{-n}\mathcal{P} \right) \\ &\geq \overline{\lim_{i\to\infty}}\, h_{\mu_i}(T, \mathcal{P}) \overset{\text{(by (ii))}}{=} \overline{\lim_{i\to\infty}}\, h_{\mu_i}(T). \end{aligned}$$

Taking the limit as $N \to \infty$ we get

$$h_\mu(T) = h_\mu(T, \mathcal{P}) = \lim_{N\to\infty} \frac{1}{N} H_\mu \left(\bigvee_{n=0}^{N-1} T^{-n}\mathcal{P} \right) \geq \overline{\lim_{i\to\infty}}\, h_{\mu_i}(T).$$

Note that we have used both (ii) and expansiveness only to establish

(ii′) $h_\nu(T) = h_\nu(T, \mathcal{P})$ for $\nu = \mu, \mu_1, \ldots$.

We could have used the following weaker condition: for every ε, there is a partition $\mathcal{P}$ satisfying (i) and

(ii″) $h_\nu(T) \leq h_\nu(T, \mathcal{P}) + \varepsilon$ for $\nu = \mu, \mu_1, \ldots$.

5.8.

We are interested in dynamical systems of the form $X = \Gamma \backslash G$ (G a connected Lie group and $\Gamma < G$ a lattice) and $T \colon x \mapsto g.x$. If G has rank ≥ 2,[22] this system will not be expansive, and furthermore in the most interesting case of $X = \mathrm{SL}(n, \mathbb{Z}) \backslash \mathrm{SL}(n, \mathbb{R})$ the space X is not compact.

Even worse, e.g., on $X = \mathrm{SL}(2, \mathbb{Z}) \backslash \mathrm{SL}(2, \mathbb{R})$ one may have a sequence of probability measures μ_i ergodic and invariant under the one parameter group $\left\{ a_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix} \right\}$ with $\underline{\lim}_{i\to\infty} h_{\mu_i}(a_\bullet) > 0$ converging weak* to a measure μ which is not a probability measure and furthermore has zero entropy[23].

However, one has the following "folklore theorem"[24] :

PROPOSITION. *Let G be a connected Lie group,* $\Gamma < G$ *a lattice, and* $H = \{a_t\}$ *a one parameter subgroup of G. Suppose that* μ_i, μ *are H-invariant* probability *measures on X with* $\mu_i \to \mu$ *in the weak* topology. Then*

$$h_\mu(a_\bullet) \geq \overline{\lim_{i\to\infty}}\, h_{\mu_i}(a_\bullet).$$

For X compact (and possibly by some clever compactification also for general X), this follows from deep (and complicated) work of Yomdin, Newhouse and Buzzi (see, e.g., (Buzzi, 1997) for more details); however Proposition 5.8 can be established quite elementarily. In order to prove this proposition, one shows that any sufficiently fine finite partition of X satisfies §5.7.(ii″).

[22] For example, $G = \mathrm{SL}(n, \mathbb{R})$ for $n \geq 3$.

[23] Strictly speaking, we define entropy only for probability measures, so one needs to rescale μ first.

[24] Which means in particular that there seems to be no good reference for it. A special case of this proposition is proved in (Einsiedler et al., 2004, Section 9). The proof of this proposition is left as an exercise to the energetic reader.

5.9.

The following example shows that this semicontinuity does not hold for a general dynamical system:

EXAMPLE. Let $S = \{1, \frac{1}{2}, \frac{1}{3}, \ldots, 0\}$, and $X = S^{\mathbb{Z}}$ (equipped with the usual Tychonoff topology). Let $\sigma\colon X \to X$ be the shift map[25].

Let μ_n be the probability measure on X obtained by taking the product of the probability measures on S giving equal probability to 0 and $1/n$, and $\delta_{\mathbf{0}}$ the probability measure supported on the fixed point $\mathbf{0} = (\ldots, 0, 0, \ldots)$ of σ. Then $\mu_n \to \delta_{\mathbf{0}}$ weak*, $h_{\mu_n}(\sigma) = \log 2$ but $h_{\delta_{\mathbf{0}}}(\sigma) = 0$.

5.10.

Let (X, d) be a compact metric space, $T\colon X \to X$ continuous[26]. Two points $x, x' \in X$ are said to be *k, ε-separated* if for some $0 \le \ell < k$ we have that $d(T^{\ell}.x, T^{\ell}.x') \ge \varepsilon$. Let $N(X, T, k, \varepsilon)$ denote the maximal cardinality of a k, ε-separated subset of X.

DEFINITION. The *topological entropy* of (X, T) is defined by

$$\begin{aligned} H(X, T, \varepsilon) &= \varlimsup_{k \to \infty} \frac{\log N(X, T, k, \varepsilon)}{k} \\ h_{\mathrm{top}}(X, T) &= \lim_{\varepsilon \to 0} H(X, T, \varepsilon). \end{aligned}$$

The topological entropy of a flow $\{a_t\}$ is defined as in §5.5 and denoted by $h_{\mathrm{top}}(X, a_\bullet)$.

5.11.

Topological entropy and the ergodic theoretic entropy are related by the *variational principle* (see, e.g., (Glasner, 2003, Theorem 17.6) or (Katok and Hasselblatt, 1995, Theorem 4.5.3))

PROPOSITION. *Let X be a compact metric space and $T : X \to X$ a homeomorphism.*[27] *Then*

$$h_{\mathrm{top}}(X, T) = \sup_{\mu} h_{\mu}(T)$$

[25] See Exercise 4.15 for a definition of the shift map.

[26] For X which is only locally compact, one can extend T to a map $\widetilde{T}$ on its one-point compactification $\tilde{X} = X \cup \{\infty\}$ fixing ∞ and define $h_{\mathrm{top}}(X, T) = h_{\mathrm{top}}(\widetilde{X}, \widetilde{T})$

[27] This proposition also easily implies the analogous statement for flows $\{a_t\}$.

where the sup *runs over all T-invariant probability measures supported on X.*

Note that when $\mu \mapsto h_\mu(T)$ is upper semicontinuous (see §5.7) the supremum is actually attained by some T-invariant measure on X.

6. Diagonalizable Actions and the Set of Exceptions to Littlewood's Conjecture

6.1.

As we have seen in §4, the action of a group H on a locally homogeneous space $X = \Gamma\backslash G$ for H generated by unipotent subgroups is quite well understood. The action of one parameter Ad-diagonalizable groups is also reasonably well understood; at least sufficiently well understood to see that there is no useful measure classification theorem in this case, since there are simply too many invariant measures (but cf. (Lindenstrauss and Schmidt, 2005, Question 1) and in a different direction §7).

Our understanding of the action of multidimensional groups H which are not generated by (Ad-)unipotents is much less satisfactory. If H contains some unipotents one can typically get quite a bit of mileage by investigating first the action of the subgroup generated by these unipotent elements (see e.g. (Margulis and Tomanov, 1996)). A typical case which is at present not well understood is the action of abelian groups H which are Ad-diagonalizable[28] over $\mathbb{R}$ with $\dim H \geq 2$, in which case one expects a "Ratner like" measure classification theorem should be true.

6.2.

Following is an explicit conjecture (essentially this is (Margulis, 2000, Conjecture 2); similar conjectures were given by Katok and Spatzier in (Katok and Spatzier, 1996) and Furstenberg (unpublished)):

CONJECTURE. *let G be a connected Lie group, $\Gamma < G$ a lattice, and $H < G$ a closed connected group generated by elements which are* Ad-*diagonalizable over $\mathbb{R}$. Let μ be a H-invariant and ergodic probability measure. Then at least one of the following holds*:

(i) *μ is homogeneous* (*cf.* §4.5)

[28] i.e., groups H whose image under the adjoint representation is diagonalizable over $\mathbb{R}$

(ii) *μ is supported on a single periodic orbit $L.x$ which has an algebraic rank one factor.*[29]

6.3.

The existence of the second, not quite algebraic, alternative in Conjecture 6.2 (§6.2.(ii)) is a complication (one of many…) we have not encountered in the theory of unipotent flows. Fortunately in some cases, in particular in the case we will focus on in this section of the full diagonal group acting on $\mathrm{SL}(n,\mathbb{Z})\backslash\mathrm{SL}(n,\mathbb{R})$, this complication can be shown not to occur, e.g. by explicitly classifying the possible H-invariant periodic orbits (not necessarily of the group H) and verifying none of them have rank one factors.[30]

6.4.

The study of such multiparameter diagonalizable actions has a long history and there are contributions by many authors. Instead of surveying this history we refer the reader to (Lindenstrauss, 2005; Einsiedler and Lindenstrauss, 2006). Rather we focus here on a specific case: $G = \mathrm{SL}(n,\mathbb{R})$, $\Gamma = \mathrm{SL}(n,\mathbb{Z})$, and $H < G$ the group of all diagonal matrices, mostly for $n = 3$, and present results from one paper (Einsiedler et al., 2004). For the remainder of this section, we set $X_n = \mathrm{SL}(n,\mathbb{Z})\backslash\mathrm{SL}(n,\mathbb{R})$.

6.5.

In this case, Conjecture 6.2 specializes to the following:

CONJECTURE. *Let H be the group of diagonal matrices in $\mathrm{SL}(n,\mathbb{R})$, $n \geq 3$. Then any H-invariant and ergodic probability measure μ on X_n is homogeneous.*

It is not hard to classify the possible homogeneous measures (see, e.g., (Lindenstrauss and Weiss, 2000)). For n prime, the situation is particularly simple: any H-invariant homogeneous measure on X_n is either the natural measure on a H-periodic orbit, or the G invariant measure m on X_n.

[29] Formally: there exists a continuous epimorphism ϕ of L onto a Lie group F such that $\phi\big(\operatorname{stab} L(x)\big)$ is closed in F and $\phi(H)$ is a one parameter subgroup of F containing no nontrivial Ad-unipotent elements.

[30] This complication does occur (Rees, 1982) when classifying invariant probability measures for certain other lattices Γ in $\mathrm{SL}(n,\mathbb{R})$ (and H the full group of diagonal matrices), and even in $\mathrm{SL}(n,\mathbb{Z})\backslash\mathrm{SL}(n,\mathbb{R})$ if one considers also infinite Radon measures.

6.6.

In (Einsiedler et al., 2004) we give the following partial result towards Conjecture 6.5:

THEOREM (Einsiedler, Katok and L.(Einsiedler et al., 2004, Theorem 1.3)). *Let H be the group of diagonal matrices as above and $n \geq 3$. Let μ be an H-invariant and ergodic probability measure on X_n. Then one of the following holds*:

(i) *μ is an H-invariant homogeneous measure which is not supported on a periodic H-orbit.*

(ii) *for every one-parameter subgroup $\{a_t\} < H$, $h_\mu(a_\bullet) = 0$.*

By the classification of H-invariant homogeneous measures alluded to in §6.5, if (i) holds μ is not compactly supported.

6.7.

Theorem 6.6 is proved by combining two techniques: a “low entropy” method developed in (Lindenstrauss, 2006b) and a “high entropy” method developed in (Einsiedler and Katok, 2003). Techniques introduced by Ratner in her study of horocycle flows in (Ratner, 1982) and subsequent papers are used in the former method. Ratner’s measure classification theorem (§4.6) is also used in the proof.

As in §4, the proof of Theorem 6.6 is beyond the scope of this paper; some hints on these methods can be found in (Lindenstrauss, 2005), but the reader who wants study the proof should consult (Einsiedler and Katok, 2003; Lindenstrauss, 2006b; Einsiedler et al., 2004).

6.8.

In §4 the fact that there were many invariant measures, even though they were explicitly given and came from countably many nice families had caused considerable difficulties when we tried to actually use this measure classification. One would think that the partial measure classification given in Theorem 6.6 would be even more difficult to use. Fortunately, this is not the case, and the key is the semicontinuity of entropy (§5.8). Using the semicontinuity one can sometimes, such as in the case of arithmetic quantum unique ergodicity considered in §7, verify positive entropy of a limiting measure by other means (see also (Einsiedler et al., 2006a)), and sometimes, such as in the case of Littlewood’s conjecture considered in this section or (Einsiedler et al., 2006b) obtain partial but meaningful results which at present cannot be obtained using alternative techniques.

6.9.

The following is a well-known conjecture of Littlewood:

CONJECTURE (Littlewood (c. 1930)). *For every $u, v \in \mathbb{R}$,*

$$\underline{\lim}_{n\to\infty} n\|nu\|\|nv\| = 0, \tag{10}$$

where $\|w\| = \min_{n\in\mathbb{Z}} |w - n|$ is the distance of $w \in \mathbb{R}$ to the nearest integer.

It turns out that this conjecture would follow from Conjecture 6.5. The reduction is nontrivial and is essentially due to Cassels and Swinnerton-Dyer (Cassels and Swinnerton-Dyer, 1955), though there is no discussion of invariant measures in that paper[31].

We need the following criterion for when α, β satisfy (10):

6.10.

PROPOSITION. *$u = \alpha$, $v = \beta$ satisfy* (10) *if and only if the orbit of*

$$x_{\alpha,\beta} = \pi_\Gamma\left(\begin{pmatrix} 1 & \alpha & \beta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)$$

under the semigroup

$$H^+ = \{a(s,t) : s, t \geq 0\} \qquad a(s,t) = \begin{pmatrix} e^{s+t} & 0 & 0 \\ 0 & e^{-s} & 0 \\ 0 & 0 & e^{-t} \end{pmatrix}$$

is unbounded[32]*. Moreover, for any $\delta > 0$ there is a compact $C_\delta \subset X_3$, so that if $\underline{\lim}_{n\to\infty} n \|n\alpha\| \|n\beta\| \geq \delta$ then $H^+.x_{\alpha,\beta} \subset C_\delta$.*

[31] It is worthwhile to note that this remarkable paper appeared in 1955, many years before Conjectures 6.2 and 6.5 were made, and even before 1967 when Furstenberg made his related discoveries about scarcity of invariant sets and measures for the maps $x \mapsto 2x \bmod 1$ and $x \mapsto 3x \bmod 1$ on $\mathbb{R}/\mathbb{Z}$! The same paper also implicitly discusses the connection between Oppenheim's conjecture and the action of $\mathrm{SO}(2,1)$ on X_3.

[32] i.e., $\overline{H^+.x_{\alpha,\beta}}$ is not compact.

6.11.

Before we prove Proposition 6.10 we need to understand better what it means for a set $E \subset X_3$ to be bounded. For this one has the following important criterion (see e.g. (Raghunathan, 1972, Chapter 10)):

PROPOSITION (Mahler's compactness criterion). *Let $n \geq 2$. A set $E \subset X_n$ is bounded if and only if there is some $\varepsilon > 0$ so that for any $x = \pi_\Gamma(g) \in X_n$ there is no vector v in the lattice spanned by the rows of g with $\|v\|_\infty < \varepsilon$.*

6.12.

We now prove Proposition 6.10.

Proof. We prove only that $H^+.x_{\alpha,\beta}$ unbounded $\Longrightarrow$ (10); the remaining assertions of this proposition follow similarly and are left as an exercise to the reader.

Let $\varepsilon \in (0, 1/2)$ be arbitrary. By Mahler's compactness criterion (§6.11), if $H^+.x_{\alpha,\beta}$ is unbounded, there is a $h \in H^+$ such that in the lattice generated by the rows of $x_{\alpha,\beta}h^{-1}$ there is a nonzero vector v with $\|v\|_\infty < \varepsilon$. This vector v is of the form

$$v = (ne^{-s-t}, (n\alpha - m)e^s, (n\beta - k)e^t)$$

where n, m, k are integers at least one of which is nonzero, and $s, t \geq 0$. Since $\|v\|_\infty < 1/2$, $n \neq 0$ and $\|n\alpha\| = (n\alpha - m)$, $\|n\beta\| = (n\beta - k)$. Without loss of generality $n > 0$ and

$$n \,\|n\alpha\|\, \|n\beta\| \leq \|v\|_\infty^3 < \varepsilon^3.$$

6.13.

We now turn to answering the following question: With the partial information given in Theorem 6.6, what information, if any, do we get regarding Littlewood's conjecture?

THEOREM ((Einsiedler et al., 2004, Theorem 1.5)). *For any $\delta > 0$, the set*

$$\Xi_\delta = \left\{ (\alpha,\beta) \in [0,1]^2 : \varliminf_{n\to\infty} n \,\|n\alpha\|\, \|n\beta\| \geq \delta \right\}$$

has zero upper box dimension[33,34].

[33] i.e.,, for every $\varepsilon > 0$, for every $0 < r < 1$, one can cover Ξ_δ by $O_{\delta,\varepsilon}(r^{-\varepsilon})$ boxes of size $r \times r$.

[34] Since (10) depends only on α, β mod 1 it is sufficient to consider only $(\alpha, \beta) \in [0, 1]^2$.

6.14.

We present a variant of the proof of this theorem given in (Einsiedler et al., 2004). The first step of the proof, which is where Theorem 6.6 is used, is an explicit sufficient criterion for a single point α, β to satisfy Littlewood's conjecture (§6.9).

Let $a_{\sigma,\tau}(t) = a(\sigma t, \tau t)$, with $a(s,t)$ as in §6.10.

PROPOSITION. *Let α, β be such that for some $\sigma, \tau \geq 0$, the topological entropy of $a_{\sigma,\tau}$ acting on*

$$\overline{\{a_{\sigma,\tau}(t).x_{\alpha,\beta} : t \in \mathbb{R}^+\}}$$

is positive. Then α, β satisfies (10).

6.15.

It will be useful for us to prove a slightly stronger result:

PROPOSITION. *Let $\sigma, \tau \geq 0$, and suppose that for $x_0 \in X_3$, the topological entropy of the action of $a_{\sigma,\tau}$ on $\overline{\{a_{\sigma,\tau}(t).x_0 : t \in \mathbb{R}^+\}}$ is positive. Then $H^+.x_0$ is unbounded.*

Note that by Proposition 6.10 the proposition above does indeed imply Proposition 6.14.

Proof. Let x_0 be as in the proposition. By the variational principle, there is a $a_{\sigma,\tau}$-invariant measure μ supported on $\overline{\{a_{\sigma,\tau}(t).x_0 : t \in \mathbb{R}^+\}}$ with $h_\mu(a_{\sigma,\tau}) > 0$.

Assume in contradiction to the proposition that $H^+.x_0$ is bounded. Define for any $S > 0$

$$\mu_S = \frac{1}{S^2} \iint_0^S a(s,t).\mu \, ds \, dt,$$

with $a(s,t).\mu$ denoting the push forward of μ under the map $x \mapsto a(s,t).x$. Since $a(s,t)$ commutes with the one parameter subgroup $a_{\sigma,\tau}$, for any $a_{\sigma,\tau}$-invariant measure μ' the entropy

$$h_{\mu'}(a_{\sigma,\tau}) = h_{a(s,t).\mu'}(a_{\sigma,\tau}).$$

If μ has the ergodic decomposition $\int \mu_\xi \, d\nu(\xi)$, the measure μ_S has ergodic decomposition $S^{-2} \iint_0^S \int a(s,t).\mu_\xi \, d\nu(\xi) \, ds \, dt$ and so by §5.3, for every S

$$h_{\mu_S}(a_{\sigma,\tau}) = h_\mu(a_{\sigma,\tau}).$$

All μ_S are supported on the compact set $\overline{H^+.x_0}$, and therefore there is a subsequence converging weak* to some compactly supported probability measure μ_∞, which will be invariant under the full group H. By semicontinuity of entropy (§5.8),

$$h_{\mu_\infty}(a_{\sigma,\tau}) \geq h_\mu(a_{\sigma,\tau}) > 0,$$

hence by Theorem 6.6 the measure μ_∞ is not compactly supported[35]—a contradiction.

6.16.

Proposition 6.14 naturally leads us to the question of the size of the set of $(\alpha,\beta) \in [0,1]^2$ for which $h_{\text{top}}(X_{\alpha,\beta}, a_{\sigma,\tau}) = 0$. This can be answered using the following general observation:

PROPOSITION. *Let X' be a metric space equipped with a continuous $\mathbb{R}$-action $(t,x) \mapsto a_t.x$. Let X'_0 be a compact $a_\bullet$-invariant*[36] *subset of X' such that for any $x \in X'_0$,*

$$h_{\text{top}}(Y_x, a_\bullet) = 0 \quad Y_x = \overline{\{a_t.x : t \in \mathbb{R}^+\}}.$$

Then $h_{\text{top}}(X'_0, a_\bullet) = 0$.

Proof. Assume in contradiction that $h_{\text{top}}(X'_0, a_\bullet) > 0$. By the variational principle (§5.11), there is some $a_\bullet$-invariant measure μ on X'_0 with $h_\mu(a_\bullet) > 0$.

By the pointwise ergodic theorem, for μ-almost every $x \in X'_0$ the measure μ is supported on Y_x. Applying the variational principle again (this time in the opposite direction) we get that

$$0 = h_{\text{top}}(Y_x, a_\bullet) \geq h_\mu(a_\bullet) > 0$$

a contradiction.

6.17.

COROLLARY. *Consider, for any compact $C \subset X_3$ the set*

$$X_C = \{x \in X_3 : H^+.x \subset C\}.$$

Then for any $\sigma, \tau \geq 0$, it holds that $h_{\text{top}}(X_C, a_{\sigma,\tau}) = 0$.

[35] Notice that a priori there is no reason to believe μ_∞ will be H-ergodic, while Theorem 6.6 deals with H-ergodic measures. So an implicit exercise to the reader is to understand why we can still deduce from $h_{\mu_\infty}(a_{\sigma,\tau}) > 0$ that μ_∞ is not compactly supported.

[36] Technical point: we only use that $a_t.X' \subset X'$ for $t \geq 0$. The variational principle (§5.11) is still applicable in this case.

Proof. By Proposition 6.15, for any $x \in X_C$ the topological entropy of $a_{\sigma,\tau}$ acting on $\overline{\{a_{\sigma,\tau}(t).x : t \in \mathbb{R}^+\}}$ is zero. The corollary now follows from Proposition 6.16.

6.18.

We are now in position to prove Theorem 6.13, or more precisely to deduce the theorem from Theorem 6.6:

Proof. To show that Ξ_δ has upper box dimension zero, we need to show, for any $\varepsilon > 0$, that for any $r \in (0,1)$ the set Ξ_δ can be covered by $O_\varepsilon(r^{-\varepsilon})$ boxes of side r, or equivalently that any r-seperated set (i.e., any set S such that for any $x, y \in S$ we have $\|x - y\|_\infty > r$) is of size $O_{\delta,\varepsilon}(r^{-\varepsilon})$.

Let C_δ be as in Proposition 6.10. Let d denote a left invariant Riemannian metric on $G = \mathrm{SL}(3,\mathbb{R})$. Then d induces a metric, also denote by d on X_3. For $a, b \in \mathbb{R}$ let

$$g_{a,b} = \begin{pmatrix} 1 & a & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Since C_δ is compact, and d induced from a left invariant Riemannian metric, there will be r_0, c_0 such that for any $x \in C_\delta$ and $|a|, |b| < r_0$

$$d(x, g_{a,b}.x) \geq c_0 \max(|a|, |b|).$$

For any $\alpha, \alpha', \beta, \beta' \in \mathbb{R}$ we have that

$$x_{\alpha,\beta} = g_{\alpha'-\alpha,\beta'-\beta}.x_{\alpha',\beta'}$$

and more generally for any t

$$a_{1,1}(t).x_{\alpha,\beta} = g_{e^{3t}(\alpha'-\alpha),e^{3t}(\beta'-\beta)}.a_{1,1}.x_{\alpha',\beta'}.$$

It follows that if $S \subset \Xi_\delta$ is r separated for $r = e^{-3t}r_0 \in (0, r_0)$ then

$$S' = \{x_{\alpha,\beta} : (\alpha,\beta) \in S\}$$

is $(t, c_0 r_0)$-separated for $a_{1,1}$ in the sense of (§5.10). By definition of C_δ and Ξ_δ, we have that (in the notations of §6.17) the set $S' \subset X_{C_\delta}$, a set which has zero topological entropy with respect to the group $a_{1,1}$. It follows that the cardinality of a maximal $(t, c_0 r_0)$-separated set in S' is at most $O_{\delta,\varepsilon}(\exp(\varepsilon t))$; hence for $r < r_0$ the cardinality of a maximal r-separated subset of Ξ_δ is $O_{\delta,\varepsilon}(r^{-\varepsilon})$.

6.19.

The alert and optimistic reader[37] may hope that there is some choice of (σ, τ), e.g $(\frac{1}{3}, \frac{2}{3})$, for which the condition of Proposition 6.14 holds for all $(\alpha,\beta) \in [0,1]^2$. To avoid trivial counterexamples of, e.g., $\alpha,\beta \in \mathbb{Q}$, we can require that for every $(\alpha,\beta) \in [0,1]^2$ either the condition of Proposition 6.14 (and hence Littlewood's conjecture) holds or $\left\{a_{\sigma,\gamma}(t).x_{\alpha,\beta} : t \in \mathbb{R}^+\right\}$ is unbounded (in which case Littlewood's conjecture follows readily from Proposition 6.10). Though I could not immediately come up with a counterexample for $(\frac{1}{3}, \frac{2}{3})$, this is extremely unlikely to be true.

We recall the following well-known conjecture of Furstenberg (which dates back to the time of (Furstenberg, 1967) but is not stated there; one place where it is explicitly stated is (Margulis, 2000, Conjecture 5)):

CONJECTURE (Furstenberg). *for any irrational $x \in \mathbb{R}/\mathbb{Z}$*

$$\frac{h_{\mathrm{top}}\left(\overline{\{2^n x : n \in \mathbb{N}\}}, \times 2\right)}{h_{\mathrm{top}}(\mathbb{R}/\mathbb{Z}, \times 2)} + \frac{h_{\mathrm{top}}\left(\overline{\{3^n x : n \in \mathbb{N}\}}, \times 3\right)}{h_{\mathrm{top}}(\mathbb{R}/\mathbb{Z}, \times 3)} \geq 1; \qquad (11)$$

In particular, the conjecture implies that for any irrational x one of the entropies in the numerators in (11) is positive. Even this is not known.

6.20.

In analogy with this conjecture, we give the following, which in view of Proposition 6.14 would imply Littlewood's conjecture (§6.9):

CONJECTURE. *Suppose $\alpha,\beta \in [0,1]^2$ is such that $\{a_{\sigma,\tau}(t).x_{\alpha,\beta} : t \in \mathbb{R}^+\}$ is bounded for both $(\sigma,\tau) = (\frac{1}{3}, \frac{2}{3})$ and $(\sigma,\tau) = (\frac{2}{3}, \frac{1}{3})$.*

Then α,β satisfy the conditions of Proposition 6.14[38] *either for $(\sigma,\tau) = (\frac{1}{3}, \frac{2}{3})$ or for $(\sigma,\tau) = (\frac{2}{3}, \frac{1}{3})$ (or both).*

One may wonder whether there are any α,β satisfying the boundedness assumptions of the conjecture. It is strongly expected that there should be many such pairs (α,β), but currently this is an open problem; indeed even the existence of one such pair is open:

6.21.

CONJECTURE ((Schmidt, 1983, p. 274)). *There is some α,β for which $\{a_{\sigma,\tau}(t).x_{\alpha,\beta} : t \in \mathbb{R}^+\}$ is bounded for both $(\sigma,\tau) = (\frac{1}{3}, \frac{2}{3})$ and $(\sigma,\tau) = (\frac{2}{3}, \frac{1}{3})$.*

[37] May he prove many theorems.

[38] i.e., the topological entropy of $a_{\sigma,\tau}$ acting on $\overline{\{a_{\sigma,\gamma}(t).x_{\alpha,\beta} : t \in \mathbb{R}^+\}}$ is positive.

The relation between this conjecture and Littlewood's conjecture is that if it is false, Littlewood's conjecture is true[39]...

7. Applications to Quantum Unique Ergodicity

7.1.

In this section we present an application of measure classification to equidistribution; but it is slightly unusual as we deal not with equidistribution of points or orbits but of *eigenfunctions*.

This equidistribution problem, a.k.a. the quantum unique ergodicity problem, is part of a much larger topic, namely the study of quantum mechanical behavior of classically chaotic systems. A basic overview of the subject was given by de Bièvre in this volume (De Bièvre, 2006), and a discussion of quantized versions of toral automorphisms by Rudnick (Rudnick, 2006).

Our discussion will be quite brief, in part because of these two contributions, and in part because the topic seems to be well covered by other sources, e.g., (Lindenstrauss, 2006a).

7.2.

Let M be a compact surfac e of constant negative curvature. Such a surface can be presented as $M = \Gamma\backslash\mathbb{H}$ with $\Gamma < \mathrm{SL}(2, \mathbb{R})$ a torsion free lattice acting on $\mathbb{H}$ by Mobius transformations. The hyperbolic plane $\mathbb{H}$ possesses a differential operator, the hyperbolic Laplacian $\Delta = y^2(\partial_x^2 + \partial_y^2)$, which is invariant under all Mobius transformations. Because of this invariance property Δ can also be viewed as a differential operator also on M.

Since M is compact, $L^2(M)$ is spanned by the Laplacian eigenfunctions, i.e., by functions ϕ_i satisfying $\Delta\phi_i = -\lambda_i\phi_i$ with $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots$; we also normalize the ϕ_i so that $\|\phi_i\|_2 = 1$.

It is natural to consider the measures $\tilde{\mu}_i$ defined by $d\tilde{\mu}_i = |\phi_i|^2\, dm$, m being the Riemannian area. For example, in quantum mechanics the eigenfunctions ϕ_i correspond to the steady states of a simple[40] particle restricted to the surface M, and given a nice measurable $A \subset M$, $\tilde{\mu}_i(A)$ is the probability of finding our particle in A (assuming we have some physical contraption which is able to measure whether or not our particle is in A).

[39] Of course the converse implication does not hold.

[40] Nonrelativistic, spinless, etc.

7.3.

Exploiting the connections between quantum and classical mechanics, Šnirel′-man, Colin de Verdière and Zelditch (Šnirel′man, 1974; Colin de Verdière, 1985; Zelditch, 1987) have shown that there is a $J \subset \mathbb{N}$ of zero density so that the sequence of measures $(\tilde{\mu}_i : i \notin J)$ converges weak* to $m(M)^{-1}m$. This phenomenon is called *quantum ergodicity*. Rudnick and Sarnak conjectured that there is no need to allow a zero density exceptional subset $J \subset \mathbb{N}$, i.e.:

CONJECTURE (Quantum Unique Ergodicity (Rudnick and Sarnak, 1994)). *For $M = \Gamma\backslash\mathbb{H}$ we have that $\tilde{\mu}_i$ converges weak* to $m(M)^{-1}m$.*

Both the quantum ergodicity theorem and the quantum unique ergodicity conjecture extend to more general M: the quantum ergodicity theorem extends to a general compact manifold M for which the geodesic flow is ergodic[41], the conjecture to compact manifolds with negative sectional curvature (not necessarily constant) in any dimension. The quantum unique ergodicity conjecture is not expected to hold in the full generality of the quantum ergodicity theorem (cf. (Donnelly, 2003)).

7.4.

So far no dynamics seem to enter; however, a key point in the proof of the quantum ergodicity theorem is that the measures $\tilde{\mu}_i$ "lift" to probability measures on the unit cotangent bundle which become increasingly invariant under the geodesic flow. Specializing to the case $M = \Gamma\backslash\mathbb{H}$ we consider, S^*M is essentially equal to[42] $X = \Gamma\backslash\,\mathrm{SL}(2,\mathbb{R})$, with the geodesic flow corresponding to the action of the diagonalizable one parameter group $a_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$. More formally, the lifting procedure allows us to define measures μ_i on X so that

(i) for any $f \in C^\infty(M)$ (viewed as a subspace of $C^\infty(X)$)

$$\left|\int f\,d\tilde{\mu}_i - \int f\,d\mu_i\right| \to 0 \quad \text{as } i \to \infty$$

(ii) for any $h \in C^\infty(X)$,

$$\int \left.\frac{dh(a_t.x)}{dt}\right|_{t=0} d\mu_i(x) \to 0 \quad \text{as } i \to \infty.$$

[41] For the Liouville measure on the unit cotangent bundle of M.

[42] More precisely, X is a double cover of the unit contangent bundle.

We call any weak* limit of the lifted measures μ_i a *quantum limit*. It follows from (i) and (ii) above that any quantum limit is a_t-invariant and that any weak* limit of the measures $\tilde{\mu}_i$ is the image of a quantum limit under the projection $\pi_{;K}: X \to M = X/K$ where K is the compact group $\mathrm{SO}(2, \mathbb{R})$.

7.5.

Without inputting any additional information on the measures $\tilde{\mu}_i$ and μ_i we cannot go further because there is no useful measure classification theorem for the action of the one parameter group a_t. Very recently, Nalini Anantharaman (Anantharaman, 2004) has been able to use the WKB approximation to obtain additional such information, proving in particular that any quantum limit has positive entropy[43] However, even with this entropy bound there is no useful measure classification.

7.6.

We vary the setting by taking Γ to be a congruence subgroup of $\mathrm{SL}(2, \mathbb{Z})$ or of certain lattices that arise from quaternionic division algebras over $\mathbb{Q}$ that are unramified over $\mathbb{R}$. The latter lattices are slightly harder to define[44] but have the advantage that X is compact.

7.7.

In both cases, for all but finitely many primes p, we have a map—the *Hecke correspondence*—T_p from X to $(p + 1)$-tuples of points of X, and a corresponding operator, also denoted by T_p, on $L^2(X)$ preserving $L^2(M)$—the *Hecke operator*. For $\Gamma = \mathrm{SL}(2, \mathbb{Z})$ these are defined in §4.25. These Hecke operators play a very important role in the spectral theory of M. In particular, the subspace of $L^2(M)$ spanned by L^2-eigenfunctions of Δ is spanned by joint eigenfunctions of Δ and all Hecke operators.[45] Let ϕ_i be such a sequence of L^2-normalized eigenfunctions of all these operators, and define using the ϕ_i measures $\tilde{\mu}_i$ and μ_i as above. Any weak* limit of the μ_i will be called an *arithmetic quantum limit*.

[43] Since quantum limits a priori need not be ergodic, this does not show that quantum limits give zero measure to a_t-periodic orbits. It does, however, show that e.g. no quantum limit gives full measure to a countable union of a_t-periodic orbits.

[44] We do not do this here; the interested reader is referred to (Lindenstrauss, 2006b) for further details.

[45] If M is not compact it is no longer true that L^2-eigenfunctions of Δ span $L^2(M)$.

7.8.

In addition to being invariant under the flow a_t, arithmetic quantum limits have the following subtle additional property (see (Lindenstrauss, 2006b, §8): T_p-recurrence.

DEFINITION. A measure μ on X is said to be *T_p-recurrent* for a fixed prime p if for any set $E \subset X$ with $\mu(E) > 0$, for μ-almost every $x \in E$ and any n, there exists an m so that

$$(T_p)^m(x) \cap \left(E - \bigcup_{k=1}^{m} (T_p)^k(x) \right) \neq \emptyset.$$

An arithmetic quantum limit is T_p-recurrent for every prime p for which T_p is defined.

7.9.

This extra recurrence assumption seems to be almost as good as having invariants under an additional one parameter group. Indeed, we conjecture the following:

CONJECTURE. *Let Γ be a congruence lattice as in* §7.6. *Let μ be an a_t-invariant probability measure on $X = \Gamma\backslash G$*[46] *which is also T_p-recurrent for a (single) prime p. Then μ is a linear combination of a_t-invariant algebraic measures.*

As in §6, this is currently known under an entropy assumption:

THEOREM ((Lindenstrauss, 2006b, Theorem 1.1)). *Let X be as in the conjecture. Let μ be an a_t-invariant probability measure on $X = \Gamma\backslash\,\mathrm{SL}(2,\mathbb{R})$ which is also*

(i) *T_p-recurrent for a (single) prime p*

(ii) *$h_{\mu_\xi}(a_\bullet) > 0$ for every a_t-invariant ergodic component μ_ξ of μ.*

then μ is the G-invariant probability measure on X.

This theorem is proved using the "low entropy" method mentioned in §6.7.

[46] Where as above $G = \mathrm{SL}(2,\mathbb{R})$.

7.10.

Condition (ii) in Theorem 7.9 was verified for quantum limits by Bourgain and the author in (Bourgain and Lindenstrauss, 2003) using combinatorial properties of Hecke points, and using all Hecke operators. No microlocal analysis was involved. Thus combining the results of (Bourgain and Lindenstrauss, 2003) and (Lindenstrauss, 2006b) one gets

THEOREM ((Lindenstrauss, 2006b, Theorem 1.4)). *Let* Γ *be a congruence lattice as in* §7.6. *Then any arithmetic quantum limit is of the form* cm, m *being the* G*-invariant measure on* X.

One would have liked to prove that c has to be $1/m(X)$, but this is currently unknown in the noncompact case (though of course this is true in the compact case).

7.11.

Theorem 7.10 is equivalent to the statement that for any $f \in C_0(X)$ with $\int f\,dm = 0$

$$\int_X f(x)\,d\mu_i(x) \to 0 \quad \text{as } i \to \infty. \tag{12}$$

In particular, if ϕ_i is a sequence of joint (say real) eigenfunctions as above

$$\int_M \phi_j(x)\phi_i(x)^2\,dx \to 0 \qquad \text{as } i \to \infty.$$

An identity of Watson (Watson, 2001) expresses this triple product in terms of L-functions; specifically for $M = \mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H}$

$$\left|\int_M \phi_j(x)\phi_i(x)^2\,dx\right|^2 = \frac{\pi^6\Lambda(1/2,\phi_i\times\phi_i\times\phi_j)}{6^6\Lambda(1,\operatorname{Sym}^2\phi_i)^2\Lambda(1,\operatorname{Sym}^2\phi_j)}$$

with Λ denoting the completed L-function. All the terms in the denominator are well understood, the numerator is much more mysterious. Currently our understanding of the numerator is not sufficient to deduce (12); but if one knew, e.g., the Riemann hypothesis for $\Lambda(s,\phi_i\times\phi_i\times\phi_j)$ one would get (12) with an optimal[47] rate of convergence. Conversely, a more quantative form of (12) would yield improved ("subconvex") estimates on $\Lambda(1/2,\phi_i\times\phi_i\times\phi_j)$.

[47] See (Luo and Sarnak, 2004, pp. 773–774).

7.12.

Anantharaman's recent results discussed in §7.5 give only information on the entropy of μ, i.e., on the average of the entropy of the ergodic components of μ but use no Hecke operators. This can be used to give weaker variants of the above theorem even if one assumes that the ϕ_i are eigenfunctions of Δ and one additional Hecke operator.

7.13.

Using the same general strategy, Silberman and Venkatesh have been able to prove a version of arithmetic quantum unique ergodicity for other $\Gamma\backslash G/K$, specifically for locally symmetric spaces arising from division algebras of prime degree. While the strategy remains the same, several new ideas are needed for this extension, in particular a new micro-local lift for higher rank groups (Silberman and Venkatesh, 2004; Silberman, 2005).

Acknowledgements

The author is grateful to Shimon Brooks, Manfred Einsiedler, Shahar Mozes and Zeev Rudnick for their comments. Discussions with Shahar and Alex Eskin were helpful in the writing of §4. For good or for bad, this paper would not have been written without the gentle insistence of the editors, Andrew Granville and Zeev Rudnick.

This research was partially supported by NSF grants DMS-0500205 and 0551823.

References

Anantharaman, N. (2004) Entropy and the localization of eigenfunctions, preprint.

Bourgain, J. and Lindenstrauss, E. (2003) Entropy of quantum limits, *Comm. Math. Phys.* **233**, 153–171.

Burger, M. (1990) Horocycle flow on geometrically finite surfaces, *Duke Math. J.* **61**, 779–803.

Burger, M. and Sarnak, P. (1991) Ramanujan duals. II, *Invent. Math.* **106**, 1–11.

Buzzi, J. (1997) Intrinsic ergodicity of smooth interval maps, *Israel J. Math.* **100**, 125–161.

Cassels, J. W. S. and Swinnerton-Dyer, H. P. F. (1955) On the product of three homogeneous linear forms and the indefinite ternary quadratic forms, *Philos. Trans. Roy. Soc. London. Ser. A.* **248**, 73–96.

Colin de Verdière, Y. (1985) Ergodicité et fonctions propres du laplacien, *Comm. Math. Phys.* **102**, 497–502.

Dani, S. G. (1978) Invariant measures of horospherical flows on noncompact homogeneous spaces, *Invent. Math.* **47**, 101–138.

Dani, S. G. and Margulis, G. A. (1989) Values of quadratic forms at primitive integral points, *Invent. Math.* **98**, 405–424.

Dani, S. G. and Margulis, G. A. (1990) Values of quadratic forms at integral points: an elementary approach, *Enseign. Math. (2)* **36**, 143–174.

Dani, S. G. and Margulis, G. A. (1993) Limit distributions of orbits of unipotent flows and values of quadratic forms, In *I. M. Gel′fand Seminar*, Vol. 16 of *Adv. Soviet Math.*, Providence, RI, Amer. Math. Soc., pp. 91–137.

De Biѐvre, S. (2006) An introduction to quantum equidistribution, in this book.

Donnelly, H. (2003) Quantum unique ergodicity, *Proc. Amer. Math. Soc.* **131**, 2945–2951.

Einsiedler, M. (2006) Ratner's theorem on $\mathrm{SL}(2,\mathbb{R})$-invariant measures, arXiv:math.DS/0603483.

Einsiedler, M. and Katok, A. (2003) Invariant measures on G/Γ for split simple Lie groups G, *Comm. Pure Appl. Math.* **56**, 1184–1221, dedicated to the memory of Jürgen K. Moser.

Einsiedler, M., Katok, A., and Lindenstrauss, E. (2004) Invariant measures and the set of exceptions to Littlewood's conjecture, *Ann. of Math. (2)*, to appear.

Einsiedler, M. and Lindenstrauss, E. (2006) Diagonal flows on locally homogeneous spaces and number theory, In *Proceedings of the International Congress of Mathematicians 2006*, to appear.

Einsiedler, M., Lindenstrauss, E., Michel, P., and Venkatesh, A. (2006a) Distribution properties of compact torus orbits. III. Duke's theorem for cubic fields, in preparation.

Einsiedler, M., Lindenstrauss, E., Michel, P., and Venkatesh, A. (2006b) Distribution properties of compact torus orbits on homogeneous spaces, in preparation.

Elkies, N. D. and McMullen, C. T. (2004) Gaps in $\sqrt{n}$ mod 1 and ergodic theory, *Duke Math. J.* **123**, 95–139.

Ellenberg, J. and Venkatesh, A. (2006) Local-global principles for representations of quadratic forms, arXiv:math.NT/0604232.

Eskin, A., Margulis, G., and Mozes, S. (1998) Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture, *Ann. of Math. (2)* **147**, 93–141.

Eskin, A., Mozes, S., and Shah, N. (1996) Unipotent flows and counting lattice points on homogeneous varieties, *Ann. of Math. (2)* **143**, 253–299.

Eskin, A. and Oh, H. (2006a) Ergodic theoretic proof of equidistribution of Hecke points, *Ergodic Theory Dynam. Systems* **26**, 163–167.

Eskin, A. and Oh, H. (2006b) Representations of integers by an invariant polynomial and unipotent flows, preprint.

Furstenberg, H. (1967) Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation, *Math. Systems Theory* **1**, 1–49.

Furstenberg, H. (1973) The unique ergodicity of the horocycle flow, In *Recent advances in topological dynamics*, Vol. 318 of *Lecture Notes in Math.*, New Haven, CO, 1972, pp. 95–115, Berlin, Springer.

Furstenberg, H. (1981) *Recurrence in ergodic theory and combinatorial number theory*, M. B. Porter Lectures, Princeton, NJ, Princeton Univ. Press.

Glasner, E. (2003) *Ergodic theory via joinings*, Vol. 101 of *Math. Surveys Monogr.*, Providence, RI, Amer. Math. Soc.

Granville, A. and Rudnick, Z. (2006) Uniform distribution, in this book.

Katok, A. and Hasselblatt, B. (1995) *Introduction to the modern theory of dynamical systems*, Vol. 54 of *Encyclopedia Math. Appl.*, Cambridge, Cambridge Univ. Press, with a supplementary chapter by A. Katok and L. Mendoza.

Katok, A. and Spatzier, R. J. (1996) Invariant measures for higher-rank hyperbolic abelian actions, *Ergodic Theory Dynam. Systems* **16**, 751–778.

Kleinbock, D., Shah, N., and Starkov, A. (2002) Dynamics of subgroup actions on homogeneous spaces of Lie groups and applications to number theory, In *Handbook of dynamical systems, Vol. 1A*, Amsterdam North-Holland, pp. 813–930.

Kleinbock, D. Y. and Margulis, G. A. (1998) Flows on homogeneous spaces and Diophantine approximation on manifolds, *Ann. of Math. (2)* **148**, 339–360.

Ledrappier, F. and Sarig, O. (2005) Invariant measures for the horocycle flow on periodic hyperbolic surfaces, *Electron. Res. Announc. Amer. Math. Soc.* **11**, 89–94.

Lindenstrauss, E. (2005) Rigidity of multiparameter actions, *Israel J. Math.* **149**, 199–226.

Lindenstrauss, E. (2006a) Arithmetic quantum unique ergodicity and adelic dynamics, In *Proceedings of Current Developments in Mathematics Conference*, Harvard, 2004, to appear.

Lindenstrauss, E. (2006b) Invariant measures and arithmetic quantum unique ergodicity, *Ann. of Math. (2)* **163**, 165–219.

Lindenstrauss, E. and Schmidt, K. (2005) Symbolic representations of nonexpansive group automorphisms, *Israel J. Math.* **149**, 227–266.

Lindenstrauss, E. and Weiss, B. (2000) Mean topological dimension, *Israel J. Math.* **115**, 1–24.

Luo, W. and Sarnak, P. (2004) Quantum variance for Hecke eigenforms, *Ann. Sci. École Norm. Sup. (4)* **37**, 769–799.

Margulis, G. (2000) Problems and conjectures in rigidity theory, In *Mathematics: frontiers and perspectives*, Providence, RI, Amer. Math. Soc., pp. 161–174.

Margulis, G. A. (1971) The action of unipotent groups in a lattice space, *Mat. Sb. (N.S.)* **86**, 552–556.

Margulis, G. A. (1989) Discrete subgroups and ergodic theory, In *Number theory, trace formulas and discrete groups*, Oslo, 1987, pp. 377–398, Boston, MA, Academic Press.

Margulis, G. A. and Tomanov, G. M. (1994) Invariant measures for actions of unipotent groups over local fields on homogeneous spaces, *Invent. Math.* **116**, 347–392.

Margulis, G. A. and Tomanov, G. M. (1996) Measure rigidity for almost linear groups and its applications, *J. Anal. Math.* **69**, 25–54.

Markloff, J. (2006) Distribution modulo one and Ratner's theorem, in this book.

Morris, D. W. (2005) *Ratner's theorems on unipotent flows*, Chicago Lectures in Math., Chicago, IL, Univ. Chicago Press.

Mozes, S. and Shah, N. (1995) On the space of ergodic invariant measures of unipotent flows, *Ergodic Theory Dynam. Systems* **15**, 149–159.

Ornstein, D. S. and Weiss, B. (1987) Entropy and isomorphism theorems for actions of amenable groups, *J. Analyse Math.* **48**, 1–141.

Raghunathan, M. S. (1972) *Discrete subgroups of Lie groups*, Vol. 68 of *Ergeb. Math. Grenzgeb.*, New York, Springer.

Ratner, M. (1982) Rigidity of horocycle flows, *Ann. of Math. (2)* **115**, 597–614.

Ratner, M. (1991a) On Raghunathan's measure conjecture, *Ann. of Math. (2)* **134**, 545–607.

Ratner, M. (1991b) Raghunathan's topological conjecture and distributions of unipotent flows, *Duke Math. J.* **63**, 235–280.

Ratner, M. (1992) Raghunathan's conjectures for $\mathrm{SL}(2,\mathbb{R})$, *Israel J. Math.* **80**, 1–31.

Ratner, M. (1995) Raghunathan's conjectures for Cartesian products of real and p-adic Lie groups, *Duke Math. J.* **77**, 275–382.

Rees, M. (1982) Some R^2-anosov flows, unpublished.

Roblin, T. (2003) *Ergodicité et équidistribution en courbure négative*, Vol. 95 of *Mém. Soc. Math. Fr. (N.S.)*, Paris, Soc. Math. France.

Rudnick, Z. (2006) The arithmetic theory of quantum maps, in this book.

Rudnick, Z. and Sarnak, P. (1994) The behaviour of eigenstates of arithmetic hyperbolic manifolds, *Comm. Math. Phys.* **161**, 195–213.

Rudolph, D. J. (1990) *Fundamentals of measurable dynamics*, Oxford Sci. Publ., New York, Oxford Univ. Press.

Schmidt, W. M. (1983) Open problems in Diophantine approximation, In *Diophantine approximations and transcendental numbers*, Vol. 31 of *Progr. Math.*, Luminy, 1982, pp. 271–287, Boston, MA, Birkhäuser Boston.

Shah, N. A. (1996) Limit distributions of expanding translates of certain orbits on homogeneous spaces, *Proc. Indian Acad. Sci. Math. Sci.* **106**, 105–125.

Silberman, L. (2005) Arithmetic quantum chaos on locally symmetric spaces, Ph.D. thesis, Princeton University.

Silberman, L. and Venkatesh, A. (2004) On quantum unique ergodicity for locally symmetric spaces. I. A micro local lift, preprint.

Šnirel′man, A. I. (1974) Ergodic properties of eigenfunctions, *Uspehi Mat. Nauk* **29**, 181–182.

Vatsal, V. (2002) Uniform distribution of Heegner points, *Invent. Math.* **148**, 1–46.

Venkatesh, A. (2006) Spectral theory of automorphic forms, a very brief introduction, in this book.

Watson, T. (2001) Rankin triple products and quantum chaos, Ph.D. thesis, Princeton University.

Zelditch, S. (1987) Uniform distribution of eigenfunctions on compact hyperbolic surfaces, *Duke Math. J.* **55**, 919–941.

AN INTRODUCTION TO QUANTUM EQUIDISTRIBUTION

S. De Bièvre
Université des Sciences et Technologies de Lille

Abstract. These notes contain crash courses on classical and quantum mechanics and on semi-classical analysis as well as a short introduction to one issue in quantum chaos: the semi-classical eigenfunction behaviour for quantum systems having an ergodic classical limit. The emphasis is on explaining the conceptual and structural similarities between the ways in which this question arises in the study of arithmetic surfaces and ergodic toral automorphisms. The text is aimed at an audience of graduate students and post-docs in number theory.

1. Introduction

These notes are loosely based on four lectures I gave during the first week of the Nato Summer School on Equidistribution in Number Theory, held at the Université de Montréal in July 2005. My task was to provide the necessary mathematics and physics background for the students to understand in which sense the topics of the second week lectures by Z. Rudnick (Rudnick, 2006) on "The arithmetic theory of quantum maps" as well as those by A. Venkatesh (Venkatesh, 2006) on the spectral analysis of the Laplace–Beltrami operator and by E. Lindenstrauss (Lindenstrauss, 2006) on quantum unique ergodicity for arithmetic surfaces can be seen as examples of a more general set of problems, referred to as "quantum chaos." In other words, I had to explain that both deal with the links between a spectral problem (the quantum side) and a Hamiltonian dynamical system (the classical side) naturally related to each other through an appropriate asymptotic analysis. For that purpose, I provided a crash course in classical mechanics and one in quantum mechanics, then gave a short introduction to semi-classical analysis, to end with an introduction to quantum maps and a proof in that context of the main equidistribution theorem in the field of quantum chaos, namely the Schnirelman theorem for quantized ergodic toral automorphisms.

Since an extensive introduction to these topics, at the beginning graduate level, can already be found in (De Bièvre, 2001), I will only briefly recall the material developed there. I will concentrate instead on developing some illustrative material (concerning symmetries in particular) that I did not have the time to develop in the actual lectures and that help to bring out the link

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 305–330.

between the two subjects alluded to above. Omitted proofs can be obtained either by a combination of matrix analysis, multivariable calculus and a little imagination, or are to be found in (De Bièvre, 2001) (or both). A recent update on what is known on equidistribution for quantum map eigenstates is available in (De Bièvre, 2005) and in the contribution of Z. Rudnick in this volume (Rudnick, 2006). Similarly, for the asymptotic behaviour of the eigenfunctions of the Laplace–Beltrami operator, the interested reader can turn to (Zelditch, 2005) and to the contribution of E. Lindenstrauss (Lindenstrauss, 2006) in this volume.

2. A Crash Course in Classical Mechanics

2.1. NEWTONIAN MECHANICS

According to Newton's second law, that you may remember from high school, "mass times acceleration equals force." In other words:

$$m\ddot{q}(t) = F(q(t)), \quad q(0) = q,\ \dot{q}(0) = v. \tag{1}$$

Here, the*force* $F\colon \mathbb{R}^n \to \mathbb{R}^n$ is given, as well as the *mass m* and the *initial data* $q, v \in \mathbb{R}^n$. The unknown in this equation is the motion of the system, namely the curve $t \in \mathbb{R} \mapsto q(t) \in \mathbb{R}^n$. In short, classical mechanics is about solving coupled non-linear second order ordinary differential equations. In most cases of interest, they are of a special type: the force is often *conservative*, meaning that $F(q) = -\nabla V(q)$ for a function $V\colon \mathbb{R}^n \to \mathbb{R}$, called *the potential*. The use of the term "conservative" is justified by the following simple result:

PROPOSITION 2.1 (Energy conservation). *Let*

$$E\colon (q, v) \in \mathbb{R}^n \times \mathbb{R}^n \to \frac{1}{2}mv^2 + V(q) \in \mathbb{R}.$$

Let $t \in \mathbb{R} \mapsto q(t) \in \mathbb{R}^n$ *be a solution to* (1)*, then, for all* $t \in \mathbb{R}$*,*

$$E(q(t), \dot{q}(t)) = E(q(0), \dot{q}(0)).$$

The function E is called the *energy* of the system (it is the sum of the *kinetic* and the *potential energy*) and the proposition states that the energy does not vary in time for a solution of (1).

It is good to keep a few examples in mind. The first one is of great historic importance and continues to attract considerable attention: it is the Kepler problem. Here $d = 3$, $V(q) = -GmM/\|q\|$, where M is the mass of the sun, and m the one of the earth, and G is the gravitational constant. Solving (1) explicitly can be done (while not trivial, it is standard...) and leads to elliptic, parabolic or hyperbolic trajectories, depending on whether the energy is strictly negative, identically zero, or strictly positive. This is a special case of a *central potential*: $V(q) = W(\|q\|)$. A second class of examples is provided by *harmonic systems*, where $V(q) = \frac{1}{2}mq^T\Omega^2 q$, with Ω^2 a positive definite n by n matrix. Now Newton's equation reads $\ddot{q} = -\Omega^2 q$. It is linear, and hence this time it *is* trivial to solve immediately:

$$q(t) = \cos\Omega t\, q + \frac{\sin\Omega t}{\Omega}\, v.$$

In general, it is of course impossible to obtain explicit solutions, and one is interested in characterizing the behaviour of the solutions, and in particular in their asymptotic properties at large times t. This will obviously depend on the type of potential one considers. For example, if $V(q) \to +\infty$ when $|q| \to +\infty$, the motion is bounded, meaning that

$$\sup_{t\in\mathbb{R}} |q(t)| \leq C < +\infty.$$

This is an easy application of energy conservation: indeed, for all t

$$V(q(t)) \leq \frac{1}{2m}\dot{q}(t)^2 + V(q(t)) = E(q(0), \dot{q}(0)). \tag{2}$$

Now, since V tends to infinity with $\|q\|$, this clearly implies (2). Such potentials are said to be *confining*.

2.2. HAMILTONIAN MECHANICS AND BEYOND

There exists an important reformulation of Newton's mechanics, referred to as Hamiltonian mechanics. Introduce the *Hamiltonian*

$$H\colon x = (q, p) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto \frac{p^2}{2m} + V(q) \in \mathbb{R}, \tag{3}$$

and observe that Newton's equation (1) is equivalent to the first-order system of differential equations called Hamilton's equations

$$\dot{q}(t) = \frac{p(t)}{m} = \frac{\partial H}{\partial p}(x(t)), \quad \dot{p} = -\nabla V(q(t)) = -\frac{\partial H}{\partial q}(x(t)), \tag{4}$$

with initial conditions $x(0) = (q, mv)$. The variable p is referred to as the *momentum* in the physics literature and the space of positions and momenta is called *phase space*. Note that the Hamiltonian is nothing but the energy expressed in terms of the position and the momentum, rather than the position and the velocity. One defines the corresponding flow $\Phi_t^H : \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ by $\Phi_t^H(x) = (q(t), p(t))$, where $q(0) = q, p(0) = p$. An obvious question that comes to mind here is: "What is the big deal?" After all, it is hard to imagine that this way of rewriting Newton's equation will shed any light on how to actually solve it. There exist at least three answers to this question. The first one is: "It's pretty! Look!" Let me compute the time rate of change of an arbitrary smooth function $f\colon \mathbb{R}^{2n} \to \mathbb{R}$ (or to $\mathbb{C}$) along a solution curve. This yields

$$\begin{aligned}\frac{d}{dt}f(q(t), p(t)) &= \partial_q f(x(t))\dot{q}(t) + \partial_p f(x(t))\dot{p}(t)\\ &= \partial_q f(x(t))\partial_p H(x(t)) - \partial_p f(x(t))\partial_q H(x(t))\\ &=: \{f, H\}(x(t)),\end{aligned}$$

where I introduced the *Poisson bracket*

$$\{\cdot,\cdot\}\colon (f, g) \in C^\infty(\mathbb{R}^{2n}) \times C^\infty(\mathbb{R}^{2n}) \mapsto \{f, g\} \in C^\infty(\mathbb{R}^{2n}),$$

with

$$\{f, g\}(x) = \partial_q f(x)\partial_p g(x) - \partial_p f(x)\partial_q g(x).$$

The reason I claim this is pretty is the following: thanks to (5)–(6) below, $C^\infty(\mathbb{R}^{2n})$ now has the structure of a Lie-algebra:

$$\{f, g\} = -\{g, f\} \qquad \text{(Anti-symmetry)} \tag{5}$$

$$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0 \qquad \text{(Jacobi identity)} \tag{6}$$

$$\{f, gh\} = \{f, g\}h + g\{f, h\} \qquad \text{(Derivation)} \tag{7}$$

If nothing else, this is certainly intriguing, and, if you are a trained mathematician of any kind, you are likely to find this pretty. But if you are nevertheless somewhat practically minded, you will be happy to know that this rewriting is also useful. To give at least one indication why (there are many others), let's have a look at constants of the motion. A *constant of the motion* is a function $f \in C^\infty(\mathbb{R}^{2n})$ which is constant along the solutions of (4), meaning that $f \circ \Phi_t^H = f$ for all $t \in \mathbb{R}$. Clearly f is a constant of the motion iff $\{f, H\} = 0$. Consequently H itself is always a constant of the motion, as we already saw. Constants of the motion are a Good Thing: the more, the merrier! Indeed, if $f_1, f_2, \ldots, f_k$ are constants of the motion, and $c \in \mathbb{R}^k$, then the flow Φ_t^H leaves their common level surface

$$\Sigma(c) = \{x \in \mathbb{R}^{2n} \mid f_j(x) = c_j, j = 1, \ldots, k\}$$

invariant. So, if they are functionally independent (meaning the Jacobian matrix $\partial_i f_j$ has rank k), the flow takes place on a $(2n - k)$-dimensional surface in $\mathbb{R}^{2n}$. This constitutes an a priori simplification of the dynamical problem, which can in this situation be thought of as a system of first order equations in $2n - k$ variables, rather than in the original $2n$ variables. Now observe that it follows immediately from the Jacobi identity that, if f and g are constants of the motion, then so is $\{f, g\}$. So, the constants of the motion make up a Lie-subalgebra of $C_0^\infty(\mathbb{R}^{2n})$ and the Poisson bracket can even be thought of as a machine for producing constants of the motion. As an example, let's look at central potentials: $V(q) = W(\|q\|)$, $q \in \mathbb{R}^3$. Then the three components $\ell_i(x) = (q \wedge p)_i$, $i = 1, \dots, 3$ of the *angular momentum* vector are constants of the motion as is readily checked through a direct computation. Note furthermore that they satisfy $\{\ell_1, \ell_2\} = \ell_3$, plus cyclic permutation. This means that the three components of the angular momentum vector form a representation of the Lie algebra so(3) of the rotation group SO(3). This is directly related to the fact that the Hamiltonian itself is invariant under the rotation group because the potential is central, as I will further discuss below.

The third answer to the question above is that the Hamiltonian formulation of classical mechanics has a number of very nice generalizations in various directions. First, *any* function $f \in C^\infty(\mathbb{R}^{2n})$ (not just the Hamiltonian) generates a flow $\Phi_t^f : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ as follows: $\Phi_t^f(x) = x(t)$ where $t \in \mathbb{R} \mapsto x(t) \in \mathbb{R}^{2n}$ solves

$$\dot{q}(t) = \frac{\partial f}{\partial p}(x(t)), \quad \dot{p} = -\frac{\partial f}{\partial q}(x(t)), \quad x(0) = x = (q, p). \tag{8}$$

I will refer to the function f as the *generator* of the flow Φ_t^f. It is clear that $\Phi_{t+t'}^f = \Phi_t^f \circ \Phi_{t'}^f$ so that the Φ_t^f define an $\mathbb{R}$-action on $\mathbb{R}^{2n}$ (meaning a group homomorphism from the additive group of reals to the diffeomorphisms of $\mathbb{R}^{2n}$), or a *Hamiltonian dynamical system*. Note that, for any $g \in C^\infty(\mathbb{R}^{2n})$, one has

$$\frac{\mathrm{d}}{\mathrm{d}t} g \circ \Phi_t^f(x) = \{g, f\}(\Phi_t^f(x)). \tag{9}$$

For further reference, let me also mention that the maps Φ_t^f are *symplectic*:

DEFINITION 2.2. A diffeomorphism Φ of $\mathbb{R}^{2n}$ is symplectic if, for all $f, g \in C^\infty(\mathbb{R}^{2n})$,

$$\{f \circ \Phi, g \circ \Phi\} = \{f, g\} \circ \Phi.$$

The group of symplectic diffeomorphisms of $\mathbb{R}^{2n}$ is denoted by $\mathrm{Diff}_{\mathrm{sympl}}(\mathbb{R}^{2n})$.

For example, on $\mathbb{R}^6$, let $f(x) = q_1 p_2 - q_2 p_1 = \ell_3(x)$. Then

$$\dot{q}_1(t) = -q_2(t), \;\; \dot{q}_2(t) = q_1(t), \;\; \dot{p}_1(t) = -p_2(t), \;\; \dot{p}_2(t) = p_1(t), \;\; \dot{q}_3 = 0 = \dot{p}_3.$$

Integrating this yields $\Phi_t^f(x) = (R_t q, R_t p)$, where $R_t = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}$ In other words, the third component of angular momentum generates rotations about the third axis (and analogously for the two other components). Using $g = H$ and $f = \ell_i$ in (9), it is now clear why the rotational invariance of H in the case the potential is central implies that each ℓ_i is a conserved quantity.

This is a general phenomenon of which we will see another example when discussing the Poincaré half plane below. A *symmetry* of a Hamiltonian H is a symplectic diffeomorphism Φ so that $H \circ \Phi = H$. If H admits a one-parameter group of symmetries of the type Φ_t^f, for some f, then, as a result of (9), f is a constant of the motion for H. In other words, if a Hamiltonian has many symmetries, it also admits many constants of the motion.

As another very simple example of a Hamiltonian flow, that will be useful in what follows, consider on $\mathbb{R}^2$ the function $f(q,p) = q^2/2$. Then $\dot{q} = 0$, $\dot{p} = -q$, so $q(t) = q$, $p(t) = p - tq$ and

$$\Phi_t^{q^2/2}(q,p) = (q, p - tq) = \begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}.$$

More generally, any function f that is a homogeneous quadratic polynomial in the variables q_i, p_i yields a linear flow.

A further and more sweeping generalization that finds its origin in Hamiltonian mechanics is symplectic geometry. Roughly, a symplectic manifold is a manifold N so that the vector space $C^\infty(N)$ is equipped with a composition law $\{\cdot,\cdot\}$, called a Poisson bracket, satisfying (5)–(7), as well as a non-degeneracy condition: $\{f,g\} = 0, \forall g \in C^\infty(N)$ implies that f is a constant. It is a non-trivial fact that, locally, there always exist coordinates (called Darboux coordinates) on N so that the Poisson bracket takes on the form it has on $\mathbb{R}^{2n}$ (implying that symplectic manifolds are always even dimensional). A simple example is the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. Write $x = (q,p) \in [0,1[^2$ and define the Poisson bracket as on $\mathbb{R}^2$. For example,

$$\{\cos q \sin p, \sin q \cos p\} = \sin^2 q \sin^2 p - \cos^2 q \cos^2 p.$$

For $n = 1$, the group $\mathrm{SL}(2,\mathbb{Z})$ (two by two matrices with determinant 1 and integer entries) acts on $\mathbb{T}^2$ by symplectic automorphisms. These maps can have rich behaviour, despite their apparent simplicity as I will discuss in more detail below. In fact, we will be picking a fixed $A \in \mathrm{SL}(2,\mathbb{Z})$ and iterate it, to obtain a $\mathbb{Z}$-action on $\mathbb{T}^2$ by symplectic transformations, or a *discrete Hamiltonian dynamical system*. This observation is the starting point for the interest in quantum maps, that I will introduce in Section 5 and that are also the subject of the contribution of Z. Rudnick in this volume (Rudnick, 2006).

A second class of important examples, in particular in connection with the lectures of E. Lindenstrauss and A. Venkatesh is provided by the *cotangent bundles* of arbitrary manifolds. Those carry a natural symplectic structure. Moreover, the geodesic flow on a Riemannian manifold can be viewed as a Hamiltonian flow on the cotangent bundle of this manifold. Rather than developing the general theory needed to understand the words appearing in the preceding sentences, an endeavour for which I have neither the space nor the inclination, I will work out the special case of the Poincaré half plane below, which is the one of relevance in the present volume.

For extensive introductions to classical mechanics, including Hamiltonian dynamics and symplectic geometry, I refer to (Arnold, 1973; Abraham and Marsden, 1978).

2.3. CLASSICAL MECHANICS ON THE POINCARÉ HALF PLANE

The Poincaré half plane is a Riemannian manifold, which is a manifold M such that at each $q \in M$, the tangent space at q is equipped with a Euclidean structure $g(q)$. The function $q \to g(q)$ is called the Riemannian structure of M. This allows one to define notions such as the length of a curve, and of a geodesic, which is a path of shortest distance between two points of the manifold. But there is no need to know Riemannian geometry to understand the basic facts about the Poincaré half plane. A few elementary and intuitive ideas from the theory of two-dimensional surfaces in $\mathbb{R}^3$ suffice largely. In fact, the Poincaré half plane can (locally) be identified with a surface of revolution, as I will explain below.

First, to understand how geodesic motion shows up in mechanics problems, let's have a look at a particle of mass m constrained to move on a sphere of radius a. One can think of the particle as being attached with a rigid rod of length a to a fixed point, taken to be the origin O. I will assume no other forces act on the particle than the pull of the rod, which is there to keep it from flying off the sphere and which acts radially. I will in particular ignore the gravitational pull on the particle, which is a reasonable thing to do when the particle moves fast (or if the rod's other end is attached to a spaceship drifting through intergalactic space). The total force acting on the particle being radial, Newton's second law implies angular momentum $\ell = mq \wedge v$ (Here the $\wedge$ stands for the vector product, not the GCD...) is conserved, as is readily seen, and consequently, the particle motion takes place in a plane through the origin: indeed, $q(t)$ is now for all times t perpendicular to the fixed vector ℓ. But the intersection of a plane through O with a sphere is, by definition, a great circle. Since moreover the force has no tangential components along the sphere, the particle's speed is constant. The particle therefore moves with constant angular speed along a great circle. Now, what is special about great circles? Well, they are precisely the geodesics or paths of shortest distance on the sphere: given two points P and B on a sphere, the path of shortest distance from P to B is the segment of the great circle obtained by intersecting the plane containing OB and OC with the sphere. This is why a flight from Paris (latitude 48.50N, longitude 2.20E) to Beijing (latitude 39.55N, longitude 116.20E) takes you over Novosibirsk (55.04N, 82.54E).

This situation generalizes to arbitrary surfaces in $\mathbb{R}^3$: if a particle is constrained to move on a surface, and the only forces that act on it are perpendicular to the surface, then one can show that Newton's second law implies it moves with constant speed along a geodesic of the surface (This takes some work to show). An interesting special case is the one of a surface of revolution $r = f(x_3)$, where (r, θ) are the polar coordinates in the x_1x_2-plane and $f\colon]\alpha, +\infty[\subset \mathbb{R}^+ \to \mathbb{R}^+$ is a nice smooth function ($0 \leq \alpha \leq 1$). Using $(\theta, x_3) \in [0, 2\pi[\times]\alpha, +\infty[$ as coordinates on the surface, and designating by v_θ, v_3 the corresponding components of a general tangent vector (check your multivariable calculus), the kinetic energy of a particle can now be written

$$E(x_3, \theta, v_3, v_\theta) = \frac{1}{2}m\left(f(x_3)^2 v_\theta^2 + (1 + f'(x_3)^2)v_3^2\right).$$

Let me introduce the new variable

$$s(x_3) = \int_1^{x_3} \sqrt{1 + f'(\zeta)^2}\mathrm{d}\zeta$$

which is the distance along the surface from the point $(x_3 = 1, \theta)$ to the point (x_3, θ) provided $x_3 \geq 1$ and minus that distance otherwise. It is convenient to use s as a coordinate instead of x_3. Note that s is a strictly growing function of x_3. In terms of the new coordinates (s, θ), the energy becomes

$$E(\theta, s, v_\theta, v_s) = \frac{1}{2}m\left(\tilde{g}(s)^2 v_\theta^2 + v_s^2\right),$$

where I introduced the function $\tilde{g}$ through the relation $\tilde{g}(s(x_3)) = f(x_3)$.

In general, given a smooth surface in $\mathbb{R}^3$, one can always introduce local coordinates $q = (q_1, q_2)$ on it that run through an open set of $\mathbb{R}^2$. Correspondingly, any tangent vector v to the surface at the point q has two components $v_1, v_2 \in \mathbb{R}$. In terms of these components, the Euclidean inner product between two tangent vectors v, w at q can be expressed in the form

$$v^T g(q) w$$

where $g(q)$ is a positive definite symmetric matrix and the T stands for transpose. For the surfaces of revolution above, with $q_1 = \theta$, $q_2 = s$, this matrix is

$$g(q) = \begin{pmatrix} \tilde{g}(q_2) & 0 \\ 0 & 1 \end{pmatrix}.$$

With this notation, the energy function can be rewritten $E(q, v) = \frac{1}{2}mv^T g(q)v$. A special case of particular relevance for us is the situation where $\tilde{g}(s) = \mathrm{e}^{-s}$. The corresponding $f(x_3)$ is easily seen to be defined only for $x_3 \geq 1$, corresponding to $s \geq 0$ (So $\alpha \geq 1$). Now, changing coordinates one last time to $x = \theta$, $y = \mathrm{e}^s$, one has, with $z = (x, y)$, $v = (v_x, v_y)$

$$E(z, v) = \frac{1}{2}m\left(\frac{v_x^2 + v_y^2}{y^2}\right) = \frac{1}{2}mv^T g(z)v, \quad g(z) = \begin{pmatrix} y^{-2} & 0 \\ 0 & y^{-2} \end{pmatrix}.$$

Now there is nothing to prevent me from extending the matrix valued function g and hence E to a function on $\mathbb{H} \times \mathbb{R}^2$, where $\mathbb{H} = \mathbb{R} \times \mathbb{R}^+$, equipped with the Riemannian structure $g(z)$ above is now, by definition, the Poincaré half plane. The surface of revolution with $f(x_3) = \mathrm{e}^{-s(x_3)}$, $x_3 \geq 1$ is obtained by quotienting the region $y \geq 1$ of the half plane by the action of $n \in \mathbb{Z}$ given by $(x, y) \to (x + 2\pi n, y)$.

Armed with these preliminaries, let me now show that the geodesic flow on $\mathbb{H}$, described in A. Venkatesh's lectures, can be viewed as a Hamiltonian flow. For that purpose, consider the following Hamiltonian function H on phase space $\mathbb{H} \times \mathbb{R}^2$:

$$H\colon (z, p) \in \mathbb{H} \times \mathbb{R}^2 \mapsto \frac{1}{2}y^2(p_x^2 + p_y^2) = \frac{1}{2}p^T g(z)^{-1}p. \tag{10}$$

The corresponding Hamilton equations of motion read

$$\dot{x} = y^2 p_x, \quad \dot{y} = y^2 p_y, \quad \dot{p}_x = 0, \quad \dot{p}_y = -y(p_x^2 + p_y^2).$$

For further reference, remark that, taking second derivatives leads to

$$\ddot{x} = 2\dot{y}y^{-1}\dot{x} = \dot{y}y^{-1}\dot{x} + \dot{x}y^{-1}\dot{y}, \quad \ddot{y} = -y^{-1}\dot{x}^2 + \dot{y}y^{-1}\dot{y}. \tag{11}$$

In view of the first two Hamilton equations above, the Hamiltonian is again nothing but the energy expressed in terms of the momentum p rather than the velocity (I put $m = 1$): $p = g(z)\dot{z}$. Remark also that the relation between the momentum p and the velocity is now position dependent, unlike what happened in (4). I claim that, given any solution $(z(t), p(t))$ of these equations, the curve $t \to z(t)$ is a geodesic on the Poincaré half plane. Since the latter are (Euclidean) half circles centered on the x axis, proving this is equivalent to showing there exists $c \in \mathbb{R}$ and $a > 0$ so that, for all t, $(x(t) - c)^2 + y(t)^2 = a^2$

or, equivalently, that $(x(t)-c)\dot{x}(t)+y(t)\dot{y}(t) = 0$. Solving for c, and expressing the result in terms of $p_x(t)$, $p_y(t)$ yields:

$$c = x(t) + y(t)\frac{p_y(t)}{p_x(t)}.$$

So the curve $t \mapsto z(t)$ is a geodesic if and only if the function

$$C\colon (z, p) \in \mathbb{H} \times \mathbb{R}^2 \mapsto x + y\frac{p_y}{p_x} \in \mathbb{R}$$

is a constant of the motion. But this is clearly the case since, as is readily checked, $\{C, H\} = 0$, so that the result follows.

Note that we found two constants of the motion, C and p_x. Since their Poisson bracket $\{C, p_x\} = 1$ is a constant, we don't find a third functionally independent constant of the motion this way. On the other hand, since we know for example from A. Venkatesh's lecture that the Poincaré half plane admits the three dimensional $\mathrm{PSL}(2,\mathbb{R})$ as group of isometries, and since isometries map geodesics into geodesics, we strongly suspect that the Hamiltonian ought to admit three one-parameter groups of symmetries and therefore have three functionally independent constants of the motion. To complete the Hamiltonian description of the geodesic flow, this is what I will now prove.

I first need some preliminaries. Any diffeomorphism φ of $\mathbb{H}$ maps a curve $\gamma\colon I \subset \mathbb{R} \to \mathbb{H}$ to a curve $\varphi \circ \gamma$ and hence the tangent vector $\dot{\gamma}(t)$ at $\gamma(t)$ to the tangent vector $D\varphi(\gamma(t)) \cdot \dot{\gamma}(t)$. Consequently, it induces a map

$$\varphi_*\colon (z, v) \in \mathbb{H} \times \mathbb{R}^2 \mapsto (\varphi(z), D\varphi(z) \cdot v) \in \mathbb{H} \times \mathbb{R}^2.$$

The diffeomorphism φ is said to be an *isometry* of the Poincaré half plane if it preserves angles between vectors and lengths of vectors, meaning that the map $D\varphi(z)$, which maps the tangent space at z to the one at $\varphi(z)$, maps the Euclidean structure $g(z)$ at z to $g(\varphi(z))$ at $\varphi(z)$:

$$D\varphi(z)^T\, g(\varphi(z))\, D\varphi(z) = g(z). \tag{12}$$

Consequently, if φ is an isometry, then $E \circ \varphi_* = E$. Now, since Hamilton's equations identify a tangent vector v with a momentum vector p via $p = g(z)v$, φ also induces a map on phase space given by

$$\varphi^*\colon (z, p) \in \mathbb{H} \times \mathbb{R}^2 \mapsto (\varphi(z), g(\varphi(z))D\varphi(z)g(z)^{-1}p) \in \mathbb{H} \times \mathbb{R}^2. \tag{13}$$

If φ is an isometry, it follows from (12) that

$$\varphi^*\colon (z, p) \in \mathbb{H} \times \mathbb{R}^2 \mapsto (\varphi(z), (D\varphi(z))^{-T} p) \in \mathbb{H} \times \mathbb{R}^2. \tag{14}$$

In that case clearly $H \circ \varphi^* = H$, proving what I promised: an isometry induces a symmetry for H. Also, as you can prove through a direct computation, φ^* is symplectic. Let me now show that to every one-parameter group of isometries φ_t corresponds a constant of the motion. Explicitly, if $X(z) := (\mathrm{d}\varphi_t/\mathrm{d}t)_{|t=0}(z)$, then $\varphi_t^* = \Phi_t^f$ with $f(z,p) := p^T X(z)$ so that f is a constant of the motion since $H \circ \Phi_t^f = H$. To see this, simply compute, using (14) and multivariable calculus

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^*(z,p)_{|t=0} &= (X(z), \frac{\mathrm{d}}{\mathrm{d}t}\left(D\varphi_{-t}(\varphi_t(z))\right)^T p)_{|t=0}\\ &= (X(z), -\partial_z X(z)^T p)\\ &= (\partial_p f(z,p), -\partial_z f(z,p)).\end{aligned}$$

We are now ready to apply this to the three one-parameter groups that generate $\mathrm{PSL}(2,\mathbb{R})$:

$$g_1(t) = \begin{pmatrix} 1 & t\\ 0 & 1\end{pmatrix} = \mathrm{e}^{t\xi_1}, \quad g_2(t) = \begin{pmatrix} 1 & 0\\ t & 1\end{pmatrix} = \mathrm{e}^{t\xi_2}, \quad g_3(t) = \begin{pmatrix} \mathrm{e}^t & 0\\ 0 & \mathrm{e}^{-t}\end{pmatrix} = \mathrm{e}^{t\xi_3}$$

with

$$\xi_1 = \begin{pmatrix} 0 & 1\\ 0 & 0\end{pmatrix}, \quad \xi_2 = \begin{pmatrix} 0 & 0\\ 1 & 0\end{pmatrix}, \quad \xi_3 = \begin{pmatrix} 1 & 0\\ 0 & -1\end{pmatrix}$$

so that

$$[\xi_1,\xi_2] = \xi_3, \quad [\xi_2,\xi_3] = 2\xi_2, \quad [\xi_3,\xi_1] = 2\xi_1. \tag{15}$$

Now, any $g = \begin{pmatrix} a & b\\ c & d\end{pmatrix}$ acts on $z = x + iy$ via $\varphi_g z = (az+b)/(cz+d)$. Defining $X_i(z) = (\mathrm{d}/\mathrm{d}t)\varphi_{it|t=0}(z)$, it is then easily checked that these vector fields are

$$X_1(z) = (1,0), \quad X_2 = (-(x^2-y^2), -2xy), \quad X_3(z) = (2x, 2y).$$

Consequently, by what precedes, $\varphi_{it}^* = \Phi_t^{f_i}$, where

$$f_1(z,p) = p_x, \quad f_2(z,p) = -(x^2-y^2)p_x - 2xyp_y, \quad f_3(z,p) = 2xp_x + 2yp_y \tag{16}$$

are now three functionally independent constants of the motion for H, a fact that can also be checked by a simple direct computation revealing that $\{f_i, H\} = 0$. Note that $C = f_3/(2f_1)$. Further simple computations reveal that

$$\{f_1, f_2\} = f_3, \quad \{f_2, f_3\} = 2f_2, \quad \{f_3, f_1\} = 2f_1.$$

so that the linear vector space spanned by the f_i is a Lie-subalgebra of $C^\infty(\mathbb{H}\times\mathbb{R}^2)$ which is isomorphic to $\mathrm{sl}(2,\mathbb{R})$, the Lie-algebra of $\mathrm{PSL}(2,\mathbb{R})$, as is clear from (15).

None of what precedes is an accident, of course, and if you suspect there must be some general theory underlying all this, you are quite right. Very briefly, the general setting is the following. Let $(N, \{\cdot, \cdot\})$ be a symplectic manifold and G a group. A symplectic action of G on N is a group homomorphism $\phi\colon g \in G \mapsto \phi_g \in \mathrm{Diff}_{\mathrm{sympl}}(N)$ (i.e., $\phi_{g_1} \circ \phi_{g_2} = \phi_{g_1 g_2}$). Such actions are particularly interesting when they provide symmetries of a given Hamiltonian dynamical system Φ_t^H: $H \circ \phi_g = H$, for all $g \in G$. An example is given by the action of $G = \mathrm{PSL}(2, \mathbb{R})$ on $N = \mathbb{H} \times \mathbb{R}^2$ described above, with H as in (10). The action ϕ is said to be *transitive* if there are no nontrivial G-invariant subsets of N. In the previous example the action is NOT transitive since each surface $H(z, p) = E$ is $\mathrm{PSL}(2, \mathbb{R})$-invariant; the action is transitive on each such surface, though. Another example is the action of SO(3) on $\mathbb{R}^6$ that we encountered previously. In that case the action is not transitive on the 5-dimensional energy surfaces $H(x) = E$, for a given central potential V, since the orbits of the action are given by the three dimensional surfaces $q^2 = R^2$, $p^2 = B^2$, $qp = C$, for some $R, B, C \in \mathbb{R}$. The presence of symmetries in a system is always a source of simplifications: in particular, the added group theoretical structures that it provides yield tools for understanding the dynamics. This is abundantly clear from the lectures of A. Venkatesh, in particular.

2.4. TWO EXTREMES: COMPLETE INTEGRABILITY AND ERGODICITY

As I pointed out from the start, we are generally interested in understanding the behaviour of the solutions of Hamilton's equations: their global properties and possibly their asymptotic behaviour in time. Two important classes of systems are the completely integrable and the ergodic ones. Let Φ_t^H be a Hamiltonian flow on $\mathbb{R}^{2n}$. It is said to be *completely integrable* if there exist n functionally independent constants of the motion $f_1 \dots f_n$, with $\{f_i, f_j\} = 0$. Supposing they are compact, it can be proven that the level surfaces $\Sigma(c) = \{x \in \mathbb{R}^{2n} \mid f_j(x) = c_j\}$ are then *n-dimensional* tori on which the Hamiltonian flow acts as a translation flow (Liouville–Arnold) (Arnold, 1973; Abraham and Marsden, 1978). The motion in such systems is very stable, in the sense that trajectories with nearby initial conditions only drift apart very slowly (linearly in time). An example are the Hamiltonians with a confining central potential: $H(x) = p^2/(2m) + W(\|q\|)$, $q, p \in \mathbb{R}^3$. One can then take $f_1 = H$, $f_2 = \ell^2$, $f_3 = \ell_3$, for example.

A Hamiltonian flow is said to be *ergodic* if a typical trajectory explores the entire $2n-1$-*dimensional energy surface* Σ_E. More precisely, given a typical trajectory on the energy surface Σ_E, the time it spends in any subset B of Σ_E is asymptotically equal to the relative size of that set in the full energy surface:

$$\lim_{T\to\infty}\frac{|\{0\le t\le T\mid x(t)\in B\}|}{T}=\frac{|B|}{|\Sigma_E|}.$$

The motion in this case can be (but need not be) very unstable: nearby trajectories may drift apart exponentially fast. This is the case for the geodesic flow on the hyperbolic surfaces as explained by A. Venkatesh. I will give another example below.

These two situations are of course very different: the present notes will be exclusively concerned with the second case.

The geodesic flow on the Poincaré half plane, viewed as a Hamiltonian flow on the corresponding phase space, is completely integrable, as I showed above. Note however that the surfaces $H = E$, $C = c$ are not tori since they are not compact: they are made up of all semi-circular orbits centered at the same point c on the x-axis, and with the same energy. In fact, the dynamics is in that case highly unstable. Now, suppose Γ is some lattice in $\mathrm{PSL}(2,\mathbb{R})$: you can then quotient $\mathbb{H}\times\mathbb{R}^2$ by the φ_g^*, $g\in\Gamma$. Since $H\circ\varphi_g^* = H$, the Hamiltonian passes to the quotient, and since functions on the quotient can be seen as Γ-periodic functions on $\mathbb{H}\times\mathbb{R}^2$, so does the Poisson bracket. As a result, the geodesic flow on $\Gamma\backslash\mathbb{H}$ is still a Hamiltonian flow, but it is no longer completely integrable! Indeed, the functions f_i are not Γ invariant and do not pass to the quotient. In fact, the flow now becomes mixing, and hence ergodic as proven in one of A. Venkatesh's lectures. This proof requires a fair amount of advanced material and preparation (in particular the representation theory of $\mathrm{PSL}(2,\mathbb{R})$), and can not be called trivial. In contrast, the simplest unstable, mixing and hence ergodic Hamiltonian dynamical systems are discrete ones, to which I now turn.

2.5. CLASSICAL MECHANICS ON THE TORUS

Consider $A\in\mathrm{SL}(2,\mathbb{Z})$, $|\mathrm{Tr}\,A| > 2$. Then A has two real eigenvalues and eigenvectors: $Av_\pm = \mathrm{e}^{\pm\gamma_0}v_\pm$. As I already pointed out, A acts as a symplectic map on $\mathbb{T}^2=\mathbb{R}^2/\mathbb{Z}^2$, meaning

$$\{f\circ A, g\circ A\} = \{f,g\}\circ A.$$

It therefore defines a discrete Hamiltonian dynamical system by iteration. This system is *hyperbolic*, meaning that for a.e. $x, x'\in\mathbb{T}^2$, $t\in\mathbb{N}$ (not too large),

$$\mathrm{d}(x,x')\sim\epsilon \Longrightarrow \mathrm{d}(A^t x, A^t x')\sim\epsilon\mathrm{e}^{\gamma_0 t}.$$

So nearby initial conditions are exponentially quickly separated from each other by the dynamics. Note that this is also a feature of the geodesic flow on the Poincaré half plane (and on its quotients by discrete subgroups Γ). This is sometimes referred to as the “butterfly effect” and is considered a crucial feature of any chaotic system. As a result of this, the maps are *exponentially mixing*: $\forall f, g \in C^\infty(\mathbb{T}^2)$,

$$\left| \int_{\mathbb{T}^2} (f \circ A^t)(x) g(x)\, \mathrm{d}x - \int_{\mathbb{T}^2} f(x)\, \mathrm{d}x \int_{\mathbb{T}^2} g(x)\, \mathrm{d}x \right| \leq C_{A,f} \|\nabla g\|_1\, \mathrm{e}^{-\gamma_0 t}.$$

Contrary to what happens with the geodesic flow on the modular surface, this is straightforward to show with a little Fourier analysis (see (De Bièvre, 2001), for example). As a result, the map is also *ergodic*: for all $f \in C^\infty(\mathbb{T}^2)$, for almost all $x_0 \in \mathbb{T}^2$

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} f(A^t x_0) = \int_{\mathbb{T}^2} f(x) \mathrm{d}x.$$

Here are some typical examples: the first is the so-called Arnold Cat Map,

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

which belongs to the following family. For $a, b \in \mathbb{N}_*$

$$A_{a,b} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix} = \Phi_a^{p^2/2} \circ \Phi_{-b}^{q^2/2} = \begin{pmatrix} 1 + ab & a \\ b & 1 \end{pmatrix}. \tag{17}$$

Another family of examples is given by $g \in \mathbb{N}_*$,

$$A_g = \begin{pmatrix} 2g & 1 \\ 4g^2 - 1 & 2g \end{pmatrix}.$$

Note that all these A act linearly on $\mathbb{R}^2$ and pass through the quotient by $\mathbb{Z}^2$ since they have integer entries. As dynamical systems on $\mathbb{R}^2$ they are already unstable, but not ergodic. The analogy with the geodesic flow on the modular surface is therefore quite clear. In both cases one starts from an unstable Hamiltonian dynamical system on a non-compact (infinite volume) space, and then considers a compact (or at least finite volume) quotient of this space on which the dynamics still acts symplectically, but is now exponentially mixing.

2.6. SUMMING UP

From a rather abstract point of view, doing classical mechanics means studying certain (discrete or continuous) symplectic dynamical system on a symplectic manifold N, equipped with a Poisson bracket $\{\cdot, \cdot\}$. Such a system is said to be chaotic if it is exponentially unstable in a suitable sense. The geodesic flow on the modular surface and the iteration of a hyperbolic $\mathrm{SL}(2,\mathbb{Z})$ matrix on the torus are two examples of such systems. Models arising in real physical problems tend to have a rather more involved behaviour, with parts of their phase space where the motion is stable, and other parts where it is unstable (to varying degrees). So, to conclude, let me say this. As a first ingredient towards understanding the interest in and the link between equidistribution for quantum maps on the torus and for the eigenfunctions of the Laplace–Beltrami operator on congruence surfaces, it is helpful to keep the following line of thinking in mind. Suitably adapted, it applies in many other scientific endeavours as well, and can be most helpful when writing introductions to scientific papers or grant applications. It goes as follows. The dynamics of many physical systems is given by a Hamiltonian flow Φ_t^H on $\mathbb{R}^{2n}$. Since physics is obviously important, it is therefore of great value to study Hamiltonian flows Φ_t^f on arbitrary symplectic manifolds. Since in such generality, only rather soft general statements can be made, it is of interest to understand relevant and tractable examples, such as geodesic flows on Riemannian manifolds, which already display a rich variety of behaviour, even if their physical pertinence is perhaps not totally clear. But since even this is often still very hard, one can hope to get insight in various issues by studying particular such manifolds, or, simpler yet, discrete Hamiltonian dynamical systems, meaning iterations of a fixed map Φ, for example an element of $\mathrm{SL}(2,\mathbb{Z})$ on $\mathbb{T}^2$.

3. A Crash Course in Quantum Mechanics

3.1. SCHRÖDINGER'S QUANTUM MECHANICS

Quantum mechanics is a physical theory that was developed in the first three decades of the twentieth century to deal with a number of issues in atomic physics that could not be dealt with using classical mechanics. I will present the theory here in a nutshell, glossing over both mathematical and physical subtleties. For some more detail, you may consult (De Bièvre, 2001) where you will find further references if you want to get serious.

According to quantum mechanics, if one wants to study the motion of a particle in a potential $V\colon \mathbb{R}^3 \to \mathbb{R}$, one should not solve Newton's equation nor Hamilton's equation, but the Schrödinger equation:

$$\mathrm{i}\hbar\frac{\partial\psi_t}{\partial t}(y) = -\frac{\hbar^2}{2m}\Delta\psi_t(y) + V(y)\psi_t(y), \quad \psi_0 = \phi,\ \|\phi\| = 1.$$

Here $\hbar$ is a physical constant called Planck's constant ($10^{-34}\,\mathrm{kgm}^2/\mathrm{s}$) and the unknown is the function $t \in \mathbb{R} \mapsto \psi_t \in L^2(\mathbb{R}^3, \mathbb{C}; \mathrm{d}y)$. One calls ψ_t *the* *wavefunction* of the particle at time t and refers to $L^2(\mathbb{R}^3)$ as the *quantum state space* or the *Hilbert space of states*. It contains all information about the particle's state. It therefore "replaces" $(q(t), p(t))$, which played this role in classical mechanics. Now this seems like a crazy idea: how can a complex valued function be used to describe the motion of a particle? According to quantum mechanics, to extract information about the particle from the wavefunction, one needs to proceed as follows. First, the probability to find the particle at time t inside a set $B \subset \mathbb{R}^3$ is $\int_B |\psi_t|^2(y)\,\mathrm{d}y$ and the probability that its momentum falls inside some set $C \subset \mathbb{R}^3$ is $\int_C |\tilde{\psi}_t|^2(p, \mathrm{d}p$, where $\tilde{\psi}_t$ is the Fourier Transform of ψ_t:

$$\tilde{\psi}_t(p) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{\mathbb{R}^3} \mathrm{e}^{-\mathrm{i}\frac{px}{\hbar}} \psi_t(y)\,\mathrm{d}y.$$

In particular, the mean position and momentum of the particle are

$$\int_{\mathbb{R}^3} y_j |\psi_t|^2(y)\,\mathrm{d}y, \qquad \int_{\mathbb{R}^3} p_j |\tilde{\psi}_t|^2(p)\mathrm{d}p.$$

Note that this makes sense since it is readily checked that the Schrödinger equation preserves the L^2-norm, so that $|\psi_t(y)|^2$ does indeed define a probability density. By the Plancherel identity, the same is then true for $|\hat{\psi}_t(p)|^2$. Since you may still find this strange, and be sceptical as to why this would have anything to do with the motion of a particle in a potential, let's try to demystify this. Introduce the following operators on (a suitable dense subspaces of) $L^2(\mathbb{R}^3, \mathrm{d}y)$:

$$P_j\psi_t(y) = \frac{\hbar}{\mathrm{i}} \frac{\partial \psi_t}{\partial y_j}(y), \qquad Q_j\psi_t(y) = y_j\psi_t(y).$$

Then

$$\langle Q_j \rangle_t := \langle Q_j\psi_t, \psi_t \rangle = \langle \psi_t, Q_j\psi_t \rangle = \int_{\mathbb{R}^3} y_j |\psi_t|^2(y)\,\mathrm{d}y$$

and

$$\langle P_j \rangle_t := \langle P_j\psi_t, \psi_t \rangle = \langle \psi_t, P_j\psi_t \rangle = \int_{\mathbb{R}^3} p_j |\tilde{\psi}_t|^2(p)\,\mathrm{d}p$$

are the mean position and momentum. So the mean position and momentum can be written as matrix elements of certain self-adjoint operators. Here

$$\langle \phi, \psi \rangle := \int_{\mathbb{R}^3} \overline{\phi}(y)\psi(y)\,\mathrm{d}y.$$

Moreover, the "Canonical Commutation Relations" hold, namely

$$[Q_j, P_k] = \mathrm{i}\hbar\delta_{jk}.$$

Now, if one defines $\widehat{H} = P^2/(2m) + V(Q)$ one can rewrite the Schrödinger equation as

$$\mathrm{i}\hbar\partial_t\psi_t = \widehat{H}\psi_t, \quad \psi_0 = \phi.$$

Since, for all $\phi, \psi \in L^2(\mathbb{R}^3, \mathrm{d}y)$, $\langle\psi, \widehat{H}\phi\rangle = \langle\widehat{H}\psi, \phi\rangle$, you can solve it through $\psi_t = \mathrm{e}^{-\mathrm{i}/\hbar t\widehat{H}}\phi$. Consequently

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi_t, Q_j\psi_t\rangle &= \langle\partial_t\psi_t, Q_j\psi_t\rangle + \langle\psi_t, Q_j\partial_t\psi_t\rangle\\ &= \frac{1}{\mathrm{i}\hbar}\langle\psi_t, [Q_j, \widehat{H}]\psi_t\rangle = \frac{1}{m}\langle\psi_t, P_j\psi_t\rangle,\end{aligned}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi_t, P_j\psi_t\rangle = \frac{1}{\mathrm{i}\hbar}\langle\psi_t, [P_j, \widehat{H}]\psi_t\rangle = -\langle\psi_t, \nabla_j V(Q)\psi_t\rangle.$$

Those equations are called the Ehrenfest equations. They should remind you of Hamilton's formulation of Newton's equation $\dot{q} = p/m$, $\dot{p} = -\nabla V(q)$. They can be paraphrased as saying that the mean momentum equals the time change of the mean position and that the time change of the mean momentum equals the mean force. Beware however: whereas Hamilton's equations are a system of ordinary differential equations for the unknown $t \mapsto (q(t), p(t))$, the Ehrenfest equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi_t, Q_j\psi_t\rangle = \frac{1}{m}\langle\psi_t, P_j\psi_t\rangle, \quad \frac{\mathrm{d}}{\mathrm{d}t}\langle\psi_t, P_j\psi_t\rangle = -\langle\psi_t, \nabla_j V(Q)\psi_t\rangle.$$

are not in general a system for $t \mapsto (\langle Q\rangle_t, \langle P\rangle_t)$! Indeed, in general

$$\langle\psi_t, \nabla_j V(Q)\psi_t\rangle = \int_{\mathbb{R}^3} \nabla_j V(y)|\psi_t(y)|^2\,\mathrm{d}y \neq \nabla_j V(\langle\psi_t, Q_j\psi_t\rangle).$$

The mean force is not equal to the value of the force at the mean position, a fact we are used to from probability theory. A notable exception is the case of a harmonic potential $V(q) = \frac{1}{2}q^T\Omega^2 q$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi_t, Q_j\psi_t\rangle = \frac{1}{m}\langle\psi_t, P_j\psi_t\rangle, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\langle\psi_t, P_j\psi_t\rangle = -\Omega^2_{jk}\langle\psi_t, Q_k\psi_t\rangle.$$

The mean position and momentum then follow the classical trajectories of the system in phase space. In particular

$$\langle Q\rangle_t = \cos\Omega t\langle Q\rangle_0 + \frac{\sin\Omega t}{\Omega}\langle P\rangle_0.$$

The moral of this story is that quadratic Hamiltonians, which give rise to a linear classical dynamics, are particularly simple, even in quantum mechanics!

In general however, to study the time evolution of the quantum system means solving the Schrödinger equation directly. Since it is a partial differential equation, this is not an easy task.

How does one go about that task? Since $\widehat{H} = \widehat{H}(\hbar)$ is self adjoint, an obvious guess is to look for an orthonormal basis of eigenfunctions of $\widehat{H}(\hbar)$:

$$\widehat{H}(\hbar)\psi_n^\hbar = E_n(\hbar)\psi_n^\hbar, \quad \text{then} \quad \psi_t = \sum_n \mathrm{e}^{-(\mathrm{i}/\hbar)E_n(\hbar)t}\langle\psi_n^\hbar, \phi\rangle\psi_n^\hbar.$$

Such a basis exists if $V(q) \to +\infty$ as $q \to \infty$. If you have enough information about the $\psi_n^\hbar$ and the $E_n(\hbar)$, you can then hope to obtain information about

$$\psi_t = \sum_{n=0}^{\infty} \mathrm{e}^{-(\mathrm{i}/\hbar)E_n(\hbar)t}\langle\psi_n^\hbar, \phi\rangle\psi_n^\hbar.$$

Of course, unless V is quadratic and in a few other special cases, it is impossible to compute the eigenfunctions and eigenvalues explicitly. Studying the behaviour of the spectrum $E_n(\hbar)$ and of the eigenfunctions $\psi_n^\hbar$ then leads to a wealth of interesting and hard problems.

One interesting question is the asymptotic behaviour along sequences $E_n(\hbar)$ that converge to a fixed value E_c as $\hbar$ goes to zero: this is part of a field called *semi-classical analysis*, for which specific techniques have been developed and to which I turn in the next section. It turns out that the behaviour of the eigenfunctions and of the eigenvalues is in that limit determined by the properties of the Hamiltonian flow of $H(q, p) = P^2/(2m) + V(q)$ on the energy surface $H(q, p) = E$.

To sum up, let me say this. In quantum mechanics the time evolution of a system is no longer given by a Hamiltonian flow Φ_t^H on a symplectic manifold (the *classical phase space*), but by a unitary flow $U_t = \mathrm{e}^{-(\mathrm{i}/\hbar)\widehat{H}t}$ on a complex Hilbert space $\mathcal{H}$ (the *quantum state space*). Whereas the symplectic flow is generated by a function H, the unitary flow is generated by a self-adjoint operator $\widehat{H}$. This is one example of a more general analogy between quantum and classical mechanics. Here is another one. Recall that the rotations R_t about the third axis are generated by $\ell_3 = q_1p_2 - q_2p_1$. Consider now the self-adjoint operator $L_3 = Q_1P_2 - Q_2P_1$. It is child's play to check that it generates a unitary group as follows:

$$\mathrm{e}^{-\mathrm{i}tL_3}\psi(y) = \psi(R_{-t}y).$$

Similarly

$$\mathrm{e}^{-\mathrm{i}a\cdot P}\psi(y) = \psi(y - a), \quad \forall a \in \mathbb{R}^3.$$

So rotations and translations act by unitary operators on $\mathcal{H}$ and their generators are functions of the position and momentum operators in complete analogy with the situation in classical mechanics.

In the previous example, $\mathcal{H} = L^2(\mathbb{R}^3)$, but other situations arise. A typical example, beyond the ones given, is

$$\mathcal{H} = L^2(M, \mathrm{dvol}_g), \quad \widehat{H} = -\hbar^2\Delta_g.$$

Here (M, g) is a (compact) Riemannian manifold and Δ_g the Laplace-Beltrami operator on it. In this case also, the asymptotic behaviour of the eigenvalues and of the eigenfunctions is sensitive to certain statistical properties of the Hamiltonian flow generated by $H(q, p) = \frac{1}{2}p^T g(q)^{-1}p$, i.e., of the geodesic flow. This is certainly not obvious a priori, but is the subject of the lectures of E. Lindenstrauss and A. Venkatesh. In the next subsection, I will show briefly, by developing the example of the half plane somewhat, why one may expect such a link between the geodesic flow and the properties of the eigenfunctions of the Laplace–Beltrami operator.

But one can also consider simpler examples, where the quantum Hilbert space is finite-dimensional and the quantum dynamics is no longer a unitary flow, but is replaced by the iteration of a fixed unitary map (so you obtain a $\mathbb{Z}$-action, rather than an $\mathbb{R}$-action). This will be the subject of Section 5 and the contribution of Z. Rudnick in this volume. In that case, as we shall see, the semi-classical limit is one in which the dimension of the finite dimensional Hilbert spaces tends to infinity. As a side remark, let me point out that in quantum computing, quantum cryptography and quantum information theory, one almost exclusively deals with finite dimensional Hilbert spaces (the N-fold tensor product of $\mathbb{C}^2$ with itself): a computation is then a certain product of unitaries.

3.2. QUANTUM MECHANICS ON THE POINCARÉ HALF PLANE

Suppose now that, armed with the insights of the previous sections, you had to (re)invent the quantum mechanics of a particle moving on the Poincaré half plane. How would you proceed? First, you would need an appropriate Hilbert space. A reasonable choice is to take $L^2(\mathbb{H}, y^{-2}\mathrm{d}x\mathrm{d}y)$, since the Riemannian volume element on $\mathbb{H}$ is $\sqrt{\det g(z)}\,\mathrm{d}x\,\mathrm{d}y = y^{-2}\,\mathrm{d}x\,\mathrm{d}y$. But how to define the appropriate quantum mechanical Hamiltonian? Pushing the analogy with the previous section, one is tempted to introduce operators X and Y, of multiplication by x and y respectively, and $P_x = (\hbar/\mathrm{i})\partial_x$, $P_y = (\hbar/\mathrm{i})\partial_y$, and then to define, inspired by (10):

$$\widehat{H} = \frac{1}{2}Y^2(P_x^2 + P_y^2) = -\frac{1}{2}y^2(\partial_x^2 + \partial_y^2).$$

With this choice, $\widehat{H} = -\frac{1}{2}\hbar^2\Delta_{\mathbb{H}}$, where $\Delta_{\mathbb{H}}$ is the Laplace–Beltrami operator on $\mathbb{H}$. Note that, with this choice, P_y is not a self-adjoint operator, but $\widehat{H}$

is, as is readily checked. The corresponding Schrödinger equation becomes $\mathrm{i}\hbar\partial_t\psi_t = \hat{H}\psi_t$.

Of course, as before, we want to interpret $|\psi_t|^2(z)$ as a probability density so that

$$\langle X\rangle_t := \int_{\mathbb{H}} x|\psi_t|^2(z)y^{-2}\,\mathrm{d}x\,\mathrm{d}y, \quad \langle Y\rangle_t := \int_{\mathbb{H}} y|\psi_t|^2(z)y^{-2}\,\mathrm{d}x\,\mathrm{d}y,$$

are the mean values of the coordinates of the particle. A simple computation, as in the derivation of the Ehrenfest equations, shows that if ψ_t solves the above Schrödinger equation, then, first of all,

$$\langle V_x\rangle_t := \frac{\mathrm{d}}{\mathrm{d}t}\langle X\rangle_t = \langle Y^2P_x\rangle_t, \quad \langle V_y\rangle_t := \frac{\mathrm{d}}{\mathrm{d}t}\langle Y\rangle_t = \langle Y^2P_y\rangle_t.$$

Note that both "velocity operators" $V_x = Y^2P_x$ and $V_y = Y^2P_y$ are self-adjoint. Furthermore

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\langle X\rangle_t = \langle V_xY^{-1}V_y + V_yY^{-1}V_x\rangle_t = \langle 2V_yY^{-1}V_x + \mathrm{i}\hbar V_x\rangle_t,$$

and

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\langle Y\rangle_t = \left\langle -\frac{1}{Y}V_x^2 + V_yY^{-1}V_y\right\rangle_t.$$

Again, comparing these last equations to (11) it is clear that the mean acceleration of the quantum particle obeys equations that bear a striking resemblance to the geodesic equations of motion on the Poincaré half plane. This begins to explain why the properties of the unitary group $\mathrm{e}^{-(\mathrm{i}/\hbar)\widehat{H}t}$ and therefore of the eigenfunctions of $\widehat{H}$ are influenced by properties of the geodesic flow. This influence is most striking asymptotically for small values of $\hbar$, as will become clearer in the following sections.

Let me end this section by a short discussion of the symmetries of the Hamiltonian on the quantum level. Let me first point out that, with the above choice of Hilbert space, the operators

$$U(g)\psi(z) = \psi(\varphi_g^{-1}(z)), \quad g \in \mathrm{PSL}(2,\mathbb{R})$$

are unitary, as is readily checked. This means the isometries of the Poincaré half plane, which are symmetries of the classical dynamical system, are realized by unitaries in the quantum Hilbert space. This is completely analogous to the representation of the Euclidean group on $L^2(\mathbb{R}^3, \mathrm{d}y)$, briefly mentioned in the previous subsection. In addition, one readily checks that

$$U(\mathrm{e}^{t\xi_j}) = \mathrm{e}^{-\mathrm{i}tF_j},\ F_1 = P_x,\ F_2 = -(X^2-Y^2)P_x-2XYP_y,\ F_3 = 2(XP_x+YP_y),$$

which is to be compared to (16). Moreover, the F_j are constants of the motion since $[\widehat{H}, F_j] = 0$ as a consequence of $U(g)^*\widehat{H}U(g) = \widehat{H}$, for all $g \in \mathrm{SL}(2, \mathbb{R})$.

4. Two Words on Semi-Classical Analysis

It was Dirac who pointed out the amazing analogy between $\{q_j, p_k\} = \delta_{jk}$ and $1/(\mathrm{i}\hbar)[Q_j, P_k] = \delta_{jk}$. This suggests that in quantum mechanics the Lie-algebra of operators on the Hilbert space $L^2(\mathbb{R}^n)$ (under the usual commutator of operators) replaces the Lie-algebra of smooth functions on phase space $\mathbb{R}^{2n}$ that appears in Hamiltonian mechanics. In classical mechanics the observables are represented by functions on phase space, in quantum mechanics by operators on a Hilbert space, as we saw on some examples. So a natural question is whether there exists a Lie-algebra homomorphism between a suitable space of functions $\mathrm{Fn}(\mathbb{R}^{2n})$ on $\mathbb{R}^{2n}$, including, say all polynomials, and the operators on $L^2(\mathbb{R}^n)$. More precisely, does there exists a linear map

$$\mathrm{Op}\colon f \in \mathrm{Fn}(\mathbb{R}^{2n}) \subset C^\infty(\mathbb{R}^{2n}) \to \mathrm{Op}(f)\colon \mathcal{D} \subset L^2(\mathbb{R}^n) \to \mathcal{D} \subset L^2(\mathbb{R}^n),$$

such that

$$\frac{1}{\mathrm{i}\hbar}[\mathrm{Op}(f), \mathrm{Op}(g)] = \mathrm{Op}(\{f, g\}),$$

and such that $\mathrm{Op}(q_j) = Q_j$, $\mathrm{Op}(p_j) = P_j$, and $\mathrm{Op}\,\bar{f} = (\mathrm{Op}\,f)^*$? It turns out that such a map does not exist, if $\mathrm{Fn}(\mathbb{R}^{2n})$ contains the space of all polynomials (Groenewold–Van Hove) (Abraham and Marsden, 1978; De Bièvre, 2001). But a map having almost those properties does exist. It was proposed by Weyl and is called the Weyl quantization or Weyl symbol calculus. It is defined as follows. For any $f \in C^\infty(\mathbb{R}^{2n})$ (of at most polynomial growth), define its Fourier transform $\tilde{f}$ by

$$f(q, p) = \int_{\mathbb{R}^{2n}} \tilde{f}(a) e^{-(\mathrm{i}/\hbar)(a_1 p - a_2 q)} \frac{\mathrm{d}a}{(2\pi\hbar)^n}.$$

Then define the *Weyl quantization of f* by

$$\mathrm{Op}^{\mathrm{W}}(f) = \int_{\mathbb{R}^{2n}} \tilde{f}(a) \mathrm{e}^{-(\mathrm{i}/\hbar)(a_1 P - a_2 Q)} \frac{\mathrm{d}a}{(2\pi\hbar)^n}.$$

Note that this is an operator since

$$U(a) := \mathrm{e}^{-(\mathrm{i}/\hbar)(a_1 P - a_2 Q)}$$

is. Of course, to give a precise mathematical meaning to the expression for Op^{W}, one needs to say in which sense the integral converges, but I will only need this formula when f is $\mathbb{Z}^2$-periodic, in which case that is a trivial matter to which I will come back in the next section. Then

$$\frac{1}{\mathrm{i}\hbar}[\mathrm{Op}^{\mathrm{W}}(f), \mathrm{Op}^{\mathrm{W}}(g)] = \mathrm{Op}^{\mathrm{W}}(\{f, g\}) + O(\hbar),$$

and there is no error term as long as f and g are polynomials of degree at most two in the q_j and the p_j. Quite explicitly, one has for example for $f(q, p) = h(q)$, respectively $g(q, p) = k(p)$

$$\mathrm{Op}^{\mathrm{W}}(f) = h(Q) \quad \mathrm{Op}^{\mathrm{W}}(g) = k(P).$$

and

$$\mathrm{Op}^{\mathrm{W}}\, q_j p_j = \frac{1}{2}(Q_j P_j + P_j Q_j).$$

A crucial, all important property of the Weyl quantization is the so-called Egorov theorem, of which I will give the following approximate statement. For all $f, g \in C^\infty(\mathbb{R}^{2n})$ (of at most polynomial growth), one has, for all $t \in \mathbb{R}$

$$\mathrm{e}^{(\mathrm{i}/\hbar)\,\mathrm{Op}^{\mathrm{W}}(g)t}\, \mathrm{Op}^{\mathrm{W}}(f) \mathrm{e}^{-(\mathrm{i}/\hbar)\,\mathrm{Op}^{\mathrm{W}}(g)t} = \mathrm{Op}^{\mathrm{W}}(f \circ \Phi_t^g) + O_t(\hbar)$$

Moreover, if g is a quadratic function, the error term vanishes.

To see why this is useful, note that $\mathrm{Op}^{\mathrm{W}}\, H = \widehat{H}$ if $H(q, p) = p^2/(2m) + V(q)$. So, since the solution ψ_t of the Schrödinger equation $\mathrm{i}\hbar\partial_t\psi_t = \widehat{H}\psi_t$ with initial condition $\psi_0 = \phi$ is $\psi_t = \mathrm{e}^{-(\mathrm{i}/\hbar)\widehat{H}t}\phi$, we find that

$$\langle\psi_t, \mathrm{Op}^{\mathrm{W}}(f)\psi_t\rangle = \langle\phi, \mathrm{Op}^{\mathrm{W}}(f \circ \Phi_t^H)\phi\rangle + O_t(\hbar).$$

This strongly suggests that, if we know enough about the classical evolution Φ_t^H appearing in the right hand side, we can infer from it information about the quantum evolution in the left hand side, in the limit of small $\hbar$. This is indeed correct and at the core of all semi-classical analysis about which much more can be learned from (Robert, 1987; Martinez, 2002). I will illustrate one aspect of this general philosophy in the remaining section.

5. Quantum Mechanics on the Torus

Let us now turn to the situation where the classical dynamics is a $\mathbb{Z}$ action on $\mathbb{T}^2$, obtained by iterating a fixed element $A \in \mathrm{SL}(2, \mathbb{Z})$ and address the following questions: What is the quantum Hilbert space of states? And the quantization of observables? And the dynamics?

Since the system has a two-dimensional phase space, it is reasonable to expect to describe the quantum states with wavefunctions $\psi(y)$ of one variable. But since the phase space is a torus, one expects that the wavefunctions must be periodic $\psi(y-1)=\psi(y)$, as well as their Fourier transforms: $\tilde{\psi}(p-1)=\tilde{\psi}(p)$. This intuition leads to the following definition. With

$$U(a)\psi(y)=\mathrm{e}^{-(\mathrm{i}/\hbar)(a_1P-a_2Q)}\psi(y)=\mathrm{e}^{-\mathrm{i}/(2\hbar)a_1a_2}\mathrm{e}^{(\mathrm{i}/\hbar)a_2y}\psi(y-a_1),$$

define

$$\mathcal{H}_\hbar=\{\psi\in\mathcal{S}'(\mathbb{R})\mid U(1,0)\psi=\psi=U(0,1)\psi\}.$$

Now, these spaces are trivial (I mean, zero-dimensional), unless there exists a positive integer such that $2\pi\hbar N=1$. So this will be assumed to be the case from now on. The semi-classical limit $\hbar\to0$ therefore becomes $N\to+\infty$. The elements of $\mathcal{H}_\hbar$ are easily described:

$$\psi\in\mathcal{H}_\hbar\Longrightarrow\psi(y)=\frac{1}{\sqrt{N}}\sum_{\ell\in\mathbb{Z}}c_\ell\delta(y-\frac{\ell}{N});\quad c_{\ell+N}=c_\ell.$$

Introducing the vectors ($j\in\mathbb{Z}$)

$$e_j=\sqrt{\frac{1}{N}}\sum_{n\in\mathbb{Z}}\delta_{j/N+n},$$

this can be written

$$\psi=\sum_{j=1}^{N}c_je_j$$

which allows one to identify $\mathcal{H}_\hbar$ with $\mathbb{C}^N$. This will be exploited in the lectures of Z. Rudnick.

As for the quantization of observables, one simply uses the Weyl quantization introduced in the previous section. For $f\in C^\infty(\mathbb{T}^2)$, $x=(q,p)\in\mathbb{T}^2$, write

$$f(x)=\sum_{n\in\mathbb{Z}^2}f_n\mathrm{e}^{-i2\pi(n_1p-n_2q)}.$$

The Weyl quantization of f can now be written

$$\mathrm{Op}^{\mathrm{W}}f=\sum_{n\in\mathbb{Z}^2}f_n\mathrm{e}^{-\mathrm{i}2\pi(n_1P-n_2Q)}=\sum_{n\in\mathbb{Z}^2}f_nU\left(\frac{n}{N}\right):\mathcal{H}_\hbar\to\mathcal{H}_\hbar.$$

The action of the phase space translation operators $U(n/N)$ on $\mathcal{H}_\hbar$ is very simple on the above basis e_j. Again, to make the link with Z. Rudnick's lectures, let me write it out in some detail. First of all, one checks that, for $n=(n_1,n_2)\in\mathbb{Z}^2$,

$$U\left(\frac{n}{N}\right)=(-1)^{n_1n_2/N}T_1^{n_1}T_2^{n_2},$$

where $T_1 = U(1/N, 0)$, $T_2 = U(0, 1/N)$. Furthermore

$$T_1 e_j = e_{j+1}, \quad T_2 e_j = e^{i2\pi j/N} e_j.$$

Up to a global normalization, the Hilbert space structure on $\mathcal{H}_\hbar$ is uniquely determined by the requirement that T_1, T_2 act unitarily, which implies that the e_j are mutually orthogonal and all have the same norm. The normalization is fixed by choosing them to be normalized. Note that any operator commuting with both T_1 and T_2, or, equivalently, with all $U(\frac{n}{N})$, is easily seen to be a multiple of the identity.

You will find the same results in Z. Rudnick's lectures, with (of course!) a slightly different notation: what he calls $\widehat{Q}$, I have called T_2 and what he calls $\widehat{P}$, I called T_1^*.

It remains to define the quantum dynamics, which ought to be a unitary map on $\mathcal{H}_\hbar$, the so-called quantum map. Let's treat the examples in (17). Defining (following Schrödinger!)

$$M(A) = \mathrm{e}^{-\mathrm{i}/(2\hbar)aP^2} \mathrm{e}^{\mathrm{i}/(2\hbar)bQ^2},$$

it is easy to check that, provided a and b are even, for all $t \in \mathbb{Z}$,

$$M(A)\mathcal{H}_\hbar = \mathcal{H}_\hbar \quad \text{and} \quad M(A)^{-t} \operatorname{Op}^{\mathrm{W}} f \, M(A)^t - \operatorname{Op}^{\mathrm{W}}(f \circ A^t) = 0.$$

This is a simple case of the Egorov theorem, and there is no error term in $\hbar$ because the dynamics is linear. A similar construction works for all hyperbolic elements of $\mathrm{SL}(2, \mathbb{Z})$ as explained in (De Bièvre, 2001; Rudnick, 2006).

To sum up, $M(A)$ is the quantum map we wish to study. It is naturally related to the discrete Hamiltonian dynamics on $\mathbb{T}^2$ obtained by iterating A through the above version of the Egorov theorem. It acts on the N dimensional spaces $\mathcal{H}_\hbar$ and we are interested in the behaviour of its eigenfunctions and eigenvalues in the $N \to \infty$ limit:

$$M(A)\psi_j^{(N)} = \mathrm{e}^{\mathrm{i}\theta_j^{(N)}} \psi_j^{(N)}, \quad j = 1, \dots, N.$$

I now finally have all the ingredients needed to state the basic result on the eigenfunction behaviour of classically ergodic systems, the so-called Schnirelman theorem, and thereby to link these lectures to the title of the school: equidistribution.

THEOREM 5.1 ((Bouzouina and De Bièvre, 1996)). *For "almost all" sequences $\psi_N \in \mathcal{H}_\hbar$, so that $M(A)\psi_N = \mathrm{e}^{\mathrm{i}\theta_N}\psi_N$,*

$$\langle \psi_N, \operatorname{Op}^{\mathrm{W}} f \psi_N \rangle \xrightarrow{N \to +\infty} \int_{\mathbb{T}^2} f(x)\,\mathrm{d}x, \quad \forall f \in C^\infty(\mathbb{T}^2). \tag{18}$$

For a simple proof of this result, and a precise explanation of what is meant by the "allmost all" in its statement, I refer to (De Bièvre, 2001) or (Rudnick, 2006). Here I just want to explain in which sense this can be seen as an equidistribution result. For that purpose, note that, given $\psi_N \in \mathcal{H}_\hbar$, one can consider the map

$$\mu_N : f \in C^\infty(\mathbb{T}^2) \to \langle \psi, \mathrm{Op}^{\mathrm{W}} f \psi \rangle \in \mathbb{C}$$

as a distribution on the torus. They are referred to as the Wigner distributions of the ψ_N. Formally, one often writes

$$\mu_N(f) = \int_{\mathbb{T}^2} \mathrm{d}x \, W_N(x) f(x),$$

but the Wigner distributions are never functions, they are in fact sums of Dirac delta measures concentrated at the points $(m_1/2N, m_2/2N)$, $0 \leq m_1, m_2 < 2N$ on the torus.

The Schnirelman theorem can therefore be paraphrased as saying that (almost all) those Wigner distributions μ_N converge to the Lebesgue measure: in other words, they equidistribute. This is the precise meaning of the much used phrase: "the eigenfunctions equidistribute in phase space." Very formally, $W_N(x) \to 1$, but of course this is not a pointwise limit. Note that this equidistribution is a reflection of the ergodicity of the underlying classical dynamical system as will be abundantly clear from the proof of the theorem, should you read it.

An analogous theorem for arithmetic surfaces can be found in the contribution of E. Lindenstrauss in this volume. The similarities with the above theorem (which is more recent) should be obvious. If they aren't, I will have done a bad job.

The Schnirelman theorem invites an obvious interrogation. Is the "almost all" in its statement an artifact of the proof or do there exist sequences of eigenfunctions for which the Wigner functions do not converge to Lebesgue measure? It was shown in (Faure et al., 2003) that there do exist such sequences. For an overview of the situation as it is understood today, I refer to (De Bièvre, 2005) as well as to the contribution of Z. Rudnick in this volume. The analogous question for arithmetic surfaces will be addressed by E. Lindenstrauss.

Number theory has played no role in my discussion. Nevertheless, it is present in the problem at hand, and it provides tools that can in particular be used to analyze the question raised in the previous paragraph as will be made clear in the lectures of Z. Rudnick .

References

Abraham, R. and Marsden, J. (1978) *Foundations of mechanics*, Reading, MA, Benjamin-Cummings.

Arnold, V. (1973) *Mathematical methods in classical mechanics*, New York–Berlin, Springer.

Bouzouina, A. and De Bièvre, S. (1996) Equipartition of the eigenfunctions of quantized ergodic maps on the torus, *Comm. Math. Phys.* **178**, 83–105.

De Bièvre, S. (2001) Quantum chaos: a brief first visit, In *Second Summer School in Analysis and Mathematical Physics*, Vol. 289, Cuernavaca, 2000, pp. 161–218, Providence, RI, Amer. Math. Soc.

De Bièvre, S. (2005) Recent results on quantum map eigenstates, In *Proceedings of QMATH 9*, Giens, 2004, preprint.

Faure, F., Nonnenmacher, S., and De Bièvre, S. (2003) Scarred eigenstates for quantum cats of minimal periods, *Comm. Math. Phys.* **239**, 449–492.

Lindenstrauss, E. (2006) Some examples how to use Measure classification in number theory, in this book.

Martinez, A. (2002) *An introduction to semiclassical and microlocal analysis*, Universitext, New York, Springer.

Robert, D. (1987) *Autour de l'approximation semi-classique*, Vol. 68 of *Progr. Math.*, Boston, MA, Birkhäuser.

Rudnick, Z. (2006) The arithmetic theory of quantum maps, in this book.

Venkatesh, A. (2006) Spectral theory of automorphic forms: a very brief introduction, in this book.

Zelditch, S. (2005) Quantum ergodicity and mixing of eigenfunctions, arXiv:math-ph/0503026.

THE ARITHMETIC THEORY OF QUANTUM MAPS

Zeév Rudnick
Tel-Aviv University

In these lectures I describe in detail the quantization of linear symplectic maps of the torus, as a continuation of De Bievre's lectures (De Bièvre, 2006). I will then survey the problem of quantum equidistribution for this model. This model was introduced by Hannay and Berry (Hannay and Berry, 1980). It turns out that it has a rich arithmetic structure, and its study uses several ingredients in modern number theory.

1. Quantum Mechanics on the Torus

We review the basics of quantum mechanics on the torus $\mathbb{T}$, viewed as a phase space, compare De Bièvre's lectures in this volume (De Bièvre, 2006). In what follows we use the abbreviations $e(z) := e^{2\pi iz}$, $e_N(z) := e^{2\pi iz/N}$.

1.1. QUANTUM STATES

We start with a description of the Hilbert space of states of such a system. In brief, Planck's constant is restricted to be an inverse integer: $h = 1/N$, and the Hilbert space of states $\mathcal{H}_N$ is N-dimensional. The "state vectors" are distributions on the line which are periodic in both momentum and position representations: $\psi(q + 1) = \psi(q)$, $[\mathcal{F}_h\psi](p + 1) = [\mathcal{F}_h\psi](p)$, where $[\mathcal{F}_h\psi](p) = h^{-1/2} \int \psi(q)\, e(-pq/h)\, dq$. The space of such distributions is finite dimensional, of dimension precisely $N = 1/h$, and consists of periodic point-masses at the coordinates $q = Q/N$, $Q \in \mathbb{Z}$. We may then identify $\mathcal{H}_N$ with the N-dimensional vector space $L^2(\mathbb{Z}/N\mathbb{Z})$, with the inner product $\langle\cdot,\cdot\rangle$ defined by

$$\langle\phi,\psi\rangle = \frac{1}{N} \sum_{Q \bmod N} \phi(Q)\overline{\psi}(Q).$$

A. Granville and Z. Rudnick (eds.), Equidistribution in Number Theory, an Introduction, 331–342.

1.2. OBSERVABLES

Classical observables, that is real-valued functions $f \in C^\infty(\mathbb{T})$, give rise to quantum observables, that is self-adjoint operators $\mathrm{Op}_N(f)$ on $\mathcal{H}_N$. To define these, one starts with the translation operators

$$[\widehat{Q}\psi](Q) = e_N(Q)\,\psi(Q), \quad \text{position}$$

and

$$[\widehat{P}\psi](Q) = \psi(Q+1), \quad \text{momentum}$$

which may be viewed as the analogues of differentiation and multiplication (respectively) operators. In fact in terms of the usual position and momentum operators on the line $\hat{q}\psi(q) = q\psi(q)$ and $\hat{p}\psi(q) = h/(2\pi i)(d/dq)\psi(q)$, they are given by $\widehat{Q} = e(\hat{q})$, $\widehat{P} = e(\hat{p})$. In this context, Heisenberg's commutation relations read

$$\widehat{P}^a\widehat{Q}^b = \widehat{Q}^b\widehat{P}^a e_N(ab) \quad \forall a,b \in \mathbb{Z}. \tag{1}$$

More generally, mixed translation operators are defined for $n = (n_1, n_2) \in \mathbb{Z}^2$ by

$$T_N(n) = e_N\left(\frac{n_1 n_2}{2}\right)\widehat{Q}^{n_2}\widehat{P}^{n_1}.$$

These are unitary operators on $\mathcal{H}_N$, whose action on a wave-function $\psi \in \mathcal{H}_N$ is given by:

$$T_N(n)\psi(Q) = e^{(i\pi n_1 n_2)/N} e\left(\frac{n_2 Q}{N}\right)\psi(Q + n_1).$$

As follows from the commutation relation (1), we have

$$T_N(m)\,T_N(n) = e_N\left(\frac{\omega(m,n)}{2}\right)T_N(m+n) \tag{2}$$

where ω is the symplectic form $\omega(m,n) = m_1 n_2 - m_2 n_1$.

For any smooth function $f \in \mathbb{C}^\infty(\mathbb{T})$, define a *quantum observable* $\mathrm{Op}_N(f)$, called the *Weyl quantization of f*, by

$$\mathrm{Op}_N(f) = \sum_{n \in \mathbb{Z}^2} \widehat{f}(n) T_N(n)$$

where $\hat{f}(n)$ are the Fourier coefficients of f.

We list some basic properties of the quantized observables $\mathrm{Op}_N(f)$:

1. The adjoint is given by

$$\mathrm{Op}_N(f)^* = \mathrm{Op}_N(\bar{f}). \tag{3}$$

2. The composition of operators satisfies:

$$\mathrm{Op}_N(f)\,\mathrm{Op}_N(g) = \mathrm{Op}_N(fg) + O_{f,g}\Big(\frac{1}{N}\Big) \tag{4}$$

for $f, g \in C^\infty(\mathbb{T})$.

3. For any orthonormal basis of $\mathcal{H}_N$ we have

$$\frac{1}{N}\sum_{j=1}^{N}\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle = \int_{\mathbb{T}} f + O_f\Big(\frac{1}{N}\Big). \tag{5}$$

That is, the *mean* of the expectation values is asymptotic to the classical average of the observable f.

Note: There is a straightforward extension of this model to quantum mechanics on tori $\mathbb{T}^{2g}$ of even dimension.

1.3. DYNAMICS

To introduce dynamics, we consider a smooth, area-preserving (symplectic) map Φ of the torus. Iterating Φ we get a discrete dynamical system. Our primary example is that of a linear automorphism $\Phi \in \mathrm{SL}_2(\mathbb{Z})$. Then the system is ergodic and in fact mixing if Φ is hyperbolic, that is $|\operatorname{tr}\Phi| > 2$. Such a map is called a “cat map” in the physics literature.

DEFINITION 1.1. A quantization of Φ is a sequence of unitary maps U_N: $\mathcal{H}_N \to \mathcal{H}_N$ such that

$$U_N^* \,\mathrm{Op}_N(f) U_N - \mathrm{Op}_N(f \circ \Phi) \to 0, \quad N \to \infty. \tag{6}$$

The operator U_N is called the *quantum propagator*, whose iterates give the evolution of the quantum system, and we require the quantum evolution to be asymptotic to the classical evolution as $N \to \infty$ (this is an analogue of “Egorov’s theorem”). In this case we say that the map Φ is “quantizable.”

In the example of the linear map Φ, one can construct a unitary operator $U_N(\Phi)$ which satisfies an *exact* version of Egorov’s theorem:

$$U_N(\Phi)^* \,\mathrm{Op}_N(f) U_N(\Phi) = \mathrm{Op}_N(f \circ \Phi). \tag{7}$$

This will be done in §2 below.

1.4. THE STONE–VON NEUMANN THEOREM

We recall the Canonical Commutation Relations (CCRN): For $\zeta_N = e^{2\pi i/N}$,

$$\widehat{P}\widehat{Q} = \zeta_N \widehat{Q}\widehat{P} \tag{8}$$

$$\widehat{P}^N = I = \widehat{Q}^N. \tag{9}$$

LEMMA 1.2. *$\widehat{P}, \widehat{Q}$ act irreducibly on $\mathcal{H}_N$*

Proof. First note that the action of $\widehat{P}, \widehat{Q}$ on δ_j is given by

$$\widehat{Q}\delta_j = \zeta_N^j \delta_j, \quad \widehat{P}\delta_j = \delta_{j-1} \tag{10}$$

Now take a subspace $0 \neq V \subseteq \mathcal{H}_N$ preserved by $\widehat{P}, \widehat{Q}$. Then $\widehat{Q}$ has an eigenvector in V. Since the only eigenvectors of $\widehat{Q}$ are δ_j (since they have distinct eigenvalues!), there is some j so that $\delta_j \in V$. But then $\widehat{P^i}\delta_j = \delta_{j-i}$ is still in V, so as i runs through $\mathbb{Z}/N\mathbb{Z}$ we get all the vectors δ_k in V, hence $V = \mathcal{H}_N$.

THEOREM 1.3. *If Q', P' are unitary operators on $\mathcal{H}_N$ satisfying the canonical commutation relations* (8) *then there is a unitary operator U for which*

$$Q' = U^* \widehat{Q} U, \quad P' = U^* \widehat{P} U \tag{11}$$

and U is unique up to a scalar multiple.

Proof. Since $Q'^N = I$, all eigenvalues of Q' are Nth roots of unity. Let

$$V_j = \{v \in \mathcal{H}_N : Q'v = \zeta_N^j v\}$$

be the eigenspace corresponding to eigenvalue ζ_N^j (possibly $V_j = 0$). If $V_j \neq 0$, then we claim that $P'V_j \subset V_{j-1}$: Indeed, if $v \in V_j$ then

$$Q'(P'v) = \zeta_N^{-1} P'Q'v = \zeta_N^{-1} P' \zeta_N^j v = \zeta_N^{j-1}(P'v)$$

which means $P'v \in V_{j-1}$. Since P' is unitary, it is in particular one-to-one so $P' : V_j \to V_{j-1}$ is one-to-one. If $0 \neq v_j \in V_j$ is a unit vector then we find $v_{j-k} := (P')^k v_j \neq 0$ is a unit vector and lies in V_{j-k}, hence for all $i \in \mathbb{Z}/N\mathbb{Z}$ we found a unit eigenvector $v_i \in V_i$ for Q': $Q'v_i = \zeta_N^i v_i$. Now define a unitary map $U \colon \mathcal{H}_N \to \mathcal{H}_N$ by taking

$$Uv_i := \sqrt{N}\delta_i, \quad i \bmod N$$

(U is unitary since it takes one orthonormal basis to another). We have $UQ' = \widehat{Q}U$ and $UP' = \widehat{P}U$ since they have the same effect on the basis vectors v_i. Thus we have constructed the required unitary map. Uniqueness follows from Lemma 1.2 by Schur's lemma.

2. Quantizing Cat Maps

Let $\Phi \in \mathrm{SL}_2(\mathbb{Z})$ and suppose Φ satisfies the parity condition

$$\Phi = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad ab \equiv cd \equiv 0 \bmod 2. \tag{12}$$

THEOREM 2.1. *For every such Φ, there is a unitary operator $U = U_N(\Phi)$ on $\mathcal{H}_N$, unique up to a phase, so that*

$$U^* \operatorname{Op}_N(f) U = \operatorname{Op}_N(f \circ \Phi),$$

Proof. We will use the Stone–von Neumann Theorem 1.3 to construct $U_N(\Phi)$. Recall that $\widehat{P} = \operatorname{Op}_N(e_{(1,0)})$, $\widehat{Q} = \operatorname{Op}_N(e_{(0,1)})$. Define

$$P' := \operatorname{Op}_N(e_{(1,0)} \circ \Phi) = \operatorname{Op}_N(e_{(1,0)\Phi}) = T_N((1,0)\Phi), \tag{13}$$

$$Q' := \operatorname{Op}_N(e_{(0,1)} \circ \Phi) = \operatorname{Op}_N(e_{(0,1)\Phi}) = T_N((0,1)\Phi) \tag{14}$$

We will show that these satisfy the canonical commutation relations (8). Firstly, we have

$$(P')^N = T_N((1,0)\Phi)^N = T_N(N(1,0)\Phi) = T_N((Na, Nb))$$

Now

$$T_N((Na, Nb))\psi(Q) = e^{(i\pi/N)N^2ab} e_N(NbQ)\psi(Q + Na) = e^{i\pi Nab}\psi(Q)$$

which equals $\psi(Q)$ since ab is even. Thus $(P')^N = I$. Likewise we have $(Q')^N = I$ since $cd \equiv 0 \bmod 2$.

Secondly, we have by (2)

$$P'Q' = T_N((1,0)\Phi)T_N((0,1)\Phi) = e^{(i\pi/N)\omega((1,0)\Phi,(0,1)\Phi)}T_N((1,1)\Phi)$$

Now since $\Phi \in \mathrm{SL}_2(\mathbb{R})$, we have $\omega(v\Phi, w\Phi) = \omega(v, w)$ and so we find

$$\omega((1,0)\Phi, (0,1)\Phi) = \omega((1,0), (0,1)) = 1$$

and so

$$P'Q' = e^{i\pi/N}T_N((1,1)\Phi).$$

Likewise we find that

$$Q'P' = e^{-i\pi/N}T_N((1,1)\Phi)$$

and so we get $Q'P' = \zeta_N^{-1}P'Q'$. Thus we have the commutation relations (8).

We may now use Theorem 1.3 to deduce the existence of U, unique up to a scalar, so that

$$P' = U^*\widehat{P}U, \quad Q' = U^*\widehat{Q}U$$

that is that

$$T_N(n\Phi) = U^*T_N(n)U \tag{15}$$

holds for the basis vectors $n = (1,0)$ and $(0,1)$. Using the relation for the product of $T_N(n)T_N(m)$ it then follows that (15) holds for all $n \in \mathbb{Z}^2$: Indeed, $T_N((n_1,n_2)) = e^{(i\pi/N)n_1n_2}\widehat{Q}^{n_2}\widehat{P}^{n_1}$ and so

$$\begin{aligned} U^*T_N(n)U &= U^*e^{(i\pi/N)n_1n_2}T_N(0,1)^{n_2}T_N(1,0)^{n_1}U \\ &= e^{(i\pi/N)n_1n_2}T_N((0,1)\Phi)^{n_2}T_N((1,0)\Phi)^{n_1} \\ &= e^{(i\pi/N)n_1n_2}T_N(n_2(0,1)\Phi)T_N(n_1(1,0)\Phi) \\ &= e^{(i\pi/N)n_1n_2}e^{(i\pi/N)\omega(n_2(0,1)\Phi,n_1(1,0)\Phi)}T_N(n\Phi). \end{aligned}$$

Now since Φ is symplectic we have

$$e^{(i\pi/N)\omega(n_2(0,1)\Phi,n_1(1,0)\Phi)} = e^{(i\pi/N)n_1n_2\omega((0,1),(1,0))} = e^{-(i\pi/N)n_1n_2}$$

and so $U^*T_N(n)U = T_N(n\Phi)$.

Hence by linearity we have

$$\mathrm{Op}_N(f \circ \Phi) = U^* \,\mathrm{Op}_N(f)U$$

for all $f \in C^\infty(\mathbb{T}^2)$.

Note: There is a similar quantization in higher dimensions, for linear symplectic maps $\Phi \in \mathrm{Sp}_{2g}(\mathbb{Z})$.

EXAMPLES.

1. For $\Phi = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ we may take the inverse Fourier transform

$$U_N\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right)\psi(Q) = \frac{1}{\sqrt{N}}\sum_{P \bmod N}\psi(P)e_N(QP).$$

2. Consider the map $\Phi = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, that is the map of the torus given by

$$\begin{aligned} p &\mapsto p + 2q \bmod 1 \\ q &\mapsto q \end{aligned}$$

The quantization is the multiplication operator $U\psi(Q) = e_N(Q^2)\psi(Q)$.

It is easy to find eigenfunctions: Indeed $U\delta_j = e_N(j^2)\delta_j$ so the position eigenstates are also eigenstates of U. To compute the matrix elements for this basis of eigenstates, note that

$$T_N(n)\delta_j = e_N\left(\frac{n_1n_2}{2} + n_2 j\right)\delta_{j-n_1}$$

and so for the normalized eigenstates $\widehat{\delta}_j := \sqrt{N}\delta_j$ we have

$$\langle T_N(n)\widehat{\delta}_j, \widehat{\delta}_j\rangle = \begin{cases} 0, & n_1 \neq 0 \bmod N \\ e_N(\frac{n_1 n_2}{2} + n_2 j), & \text{otherwise} \end{cases}$$

Therefore as $N \to \infty$,

$$\begin{aligned} \langle \mathrm{Op}_N(f)\hat{\delta}_j, \hat{\delta}_j\rangle &= \sum_{n_1 \equiv 0 \bmod N} \hat{f}(n_1, n_2) e\Big(\frac{n_1 n_2}{2N} + \frac{n_2 j}{N}\Big) \\ &\sim \sum_{n_2} \widehat{f}(0, n_2) e\Big(\frac{n_2 j}{N}\Big) \\ &= \int_0^1 f(p, \frac{j}{N})\, dp = \int_{\mathbb{T}^2} f\, dp \otimes \delta_{j/N}(q) dq. \end{aligned}$$

Thus if $j/N \to q_0$ we get $dp \otimes \delta(q - q_0)dq$ as the corresponding quantum limit. These are delta-functions on invariant curves $q = q_0$ of the classical map.

3. Quantum Ergodicity

One of the fundamental questions in the subject of “quantum chaos” is the limiting behaviour of eigenfunctions. In particular, given an observable $f \in C^\infty(\mathbb{T}^2)$, we would like to study the possible limits of matrix elements $\langle \mathrm{Op}_N(f), \psi, \psi\rangle$ where $\psi \in \mathcal{H}_N$ are eigenfunctions of $U_N(\Phi)$. It turns out that these limits (as functionals of the observable f) have to be probability measures on the torus, invariant under the map Φ. A very general result says that at least for almost all choices of eigenfunctions, the limit must be Lebesgue measure. This is known as the quantum ergodicity theorem. In the context of ergodic geodesic flows this was formulated by Schnirelman (Schnirelman, 1974) and proved in (Zelditch, 1987; Colin de Verdière, 1985). For quantum maps, the form that quantum ergodicity assumes is the following:

THEOREM 3.1 ((Bouzouina and De Bièvre 1996; Zelditch, 1996, 1997)). *Let $\Phi \in \mathrm{SL}_2(\mathbb{Z})$ be ergodic. Then for any orthonormal basis ψ_j of $\mathcal{H}_N$ consisting of eigenfunctions of $U_N(\Phi)$, there is a subset $J(N) \subset \{1, 2, \ldots, N\}$, with $\#J(N)/N \to 1$, so that for $j \in J(N)$ we have:*

$$\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle \to \int_{\mathbb{T}} f, \quad \textit{as } N \to \infty$$

for all observables $f \in C^\infty(\mathbb{T})$.

Theorem 3.1 is a consequence, using positivity and a standard diagonalization argument, of the following estimate for the variance due to Zelditch (Zelditch, 1996).

THEOREM 3.2. *Let $\Phi \in \mathrm{SL}_2(\mathbb{Z})$ be ergodic. For any orthonormal basis ψ_j, $j = 1, \dots, N$ of of $\mathcal{H}_N$ consisting of eigenfunctions of $U_N(\Phi)$, we have*

$$\frac{1}{N}\sum_{j=1}^{N}\left|\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle - \int_{\mathbb{T}} f\right|^2 \to 0$$

for all observables $f \in C^\infty(\mathbb{T})$.

Proof. Without loss of generality, we may assume that $\int_{\mathbb{T}} f = 0$.

We fix $T \geq 1$. By Egorov's theorem (7), we have

$$\frac{1}{T}\sum_{j=1}^{T}(U_N(\Phi)^t)^* \,\mathrm{Op}_N(f) U_N(\Phi)^t = \frac{1}{T}\sum_{t=1}^{T}\mathrm{Op}_N(f \circ \Phi^t) = \mathrm{Op}_N\left(f^T\right)$$

where $f^T := (1/T)\sum_{t=1}^{T} f \circ \Phi^t$ is the ergodic average of f. Moreover, if ψ_j is an eigenfunction: $U_N(\Phi)\psi_j = e^{i\lambda_j}\psi_j$, then

$$\begin{aligned}\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle &= \langle \mathrm{Op}_N(f) U_N(\Phi)\psi_j, U_N(\Phi)\psi_j\rangle \\ &= \langle U_N(\Phi)^* \,\mathrm{Op}_N(f) U_N(\Phi)\psi_j, \psi_j\rangle \\ &= \langle \mathrm{Op}_N(f \circ \Phi)\psi_j, \psi_j\rangle.\end{aligned}$$

Consequently, if ψ_j is an eigenfunction then for all $T \geq 0$,

$$\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle = \left\langle \mathrm{Op}_N\left(f^T\right)\psi_j, \psi_j\right\rangle. \tag{16}$$

Now we look at the sum (recall that $\int_{\mathbb{T}} f = 0$)

$$S_2(f, N) := \frac{1}{N}\sum_{j=1}^{N}\left|\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle\right|^2.$$

We will show that $\lim_{N\to\infty} S_2(f, N) = 0$.

By (16), we have $S_2(f, N) = S_2(f^T, N)$ for all $T \geq 1$. By Cauchy–Schwartz, we have

$$\left|\left\langle \mathrm{Op}_N\left(f^T\right)\psi_j, \psi_j\right\rangle\right|^2 \leq \left\|\mathrm{Op}_N\left(f^T\right)\psi_j\right\|^2 \|\psi_j\|^2 = \left\langle \mathrm{Op}_N\left(f^T\right)^* \mathrm{Op}_N\left(f^T\right)\psi_j, \psi_j\right\rangle.$$

Moreover, by (3), (4),

$$\mathrm{Op}_N\left(f^T\right)^* \mathrm{Op}_N\left(f^T\right) = \mathrm{Op}_N\left(\left|f^T\right|^2\right) + O_{f,T}\left(\frac{1}{N}\right)$$

and so

$$S_2(f,N) \lesssim \frac{1}{N}\sum_{j=1}^{N}\left\langle \mathrm{Op}_N\left(\left|f^T\right|^2\right)\psi_j,\psi_j\right\rangle + O_{f,T}\left(\frac{1}{N}\right).$$

By (5) we thus find that for fixed $T \geq 1$,

$$\limsup S_2(f,N) \leq \int_{\mathbb{T}}\left|f^T\right|^2.$$

So far we have used nothing about the cat map except Egorov's theorem. Now we use the fact that it is *ergodic*, in particular the *mean ergodic theorem* holds: For $F \in L^2(\mathbb{T})$, the ergodic averages F^T converge to $\int_{\mathbb{T}} F$ in L^2. Thus we have $\int_{\mathbb{T}} |f^T|^2 \to 0$ as $T \to \infty$. Therefore given $\epsilon > 0$, we can find $T = T(f,\epsilon)$ for which $\int_{\mathbb{T}} |f^T|^2 < \epsilon$ and consequently

$$\limsup S_2(f,N) < \epsilon$$

which shows that $S_2(f,N) \to 0$ as required.

Note that the argument used nothing more than Egorov's theorem and the *ergodicity* of the map. In fact it gives the result for any symplectic ergodic (quantizable) map of the torus.

4. Quantum Unique Ergodicity

A key problem is:

PROBLEM 4.1. Is it true that *all* eigenfunctions become equidistributed as $N \to \infty$?

In the analogous situation of the quantization of the geodesic flow on a surface of negative curvature, it is conjectured (Rudnick and Sarnak, 1994) that all eigenfunctions are equidistributed. This is the Quantum Unique Ergodicity conjecture (QUE), see Lindenstrauss' lectures (Lindenstrauss, 2006). It was thus a surprise when De Bièvre, Faure and Nonnenmacher (Faure et al., 2003) discovered that this is false for cat maps. For instance, there is a (sparse) sequence of inverse Planck constants N_k and eigenfunctions $\psi_k \in \mathcal{H}_{N_k}$ for which the matrix coefficients converge to $\frac{1}{2}(\delta_0 + dx)$.

The existence of these "scars" is related to the fact that eigenspaces of $U_N(\Phi)$ may have large dimension. In fact, the *mean* degeneracy is $N/\operatorname{ord}(\Phi,N)$ where $\operatorname{ord}(\Phi,N)$ the *order* (or period) of Φ modulo N, that is the least integer $k \geq 1$ for which $\Phi^k = I \bmod N$. It can be shown that the mean degeneracy can be as large as $N/\log N$ for arbitrarily large N. It is precisely for these N's that the authors of (Faure et al., 2003) constructed "scars," that is eigenfunctions which are not uniformly distributed.

These "scars" are rare; for instance, Kurlberg–Rudnick show (Kurlberg and Rudnick, 2001) that for almost all values of N (that is the exceptional set has density zero), all eigenfunctions are uniformly distributed. On GRH, this holds for almost all *prime* values of N (Kurlberg, 2003). This has recently been improved by Bourgain (Bourgain, 2006) who showed unconditionally that the result holds for almost all primes and that the exceptional set is sparse. Bourgain's main new ingredient are his estimates on incomplete exponential sums, encountered already in Friedlander's lectures (Friedlander, 2006).

5. Arithmetic QUE

5.1. SYMMETRIES AND HECKE OPERATORS

It transpires that there is a commutative group of unitary operators on the state-space $\mathcal{H}_N$ which commute with the quantized map and therefore act on its eigenspaces. These operators were discovered by Kurlberg and the author (Kurlberg and Rudnick, 2000), where we called them "Hecke operators," in analogy with the setting of the modular surface.

To understand their origin, one needs to note that the definition of $U_N(\Phi)$ given in §2 is such that it only depends on the remainder of $\Phi \bmod 2N$. From the fact that the Egorov property (7) holds one thus gets a *projective* representation $\Phi \mapsto U_N(\Phi)$ of the subgroup of "quantizable" elements in the finite modular group $\mathrm{SL}(2, \mathbb{Z}/2N\mathbb{Z})$. It turns out that it can be made into an *ordinary* representation if we further restrict to the subgroup $\Gamma(4, 2N)$ given by $g = I \bmod 4$ for N even, $g = I \bmod 2$ for N odd. Thus for $\Phi, \Phi' \in \Gamma(4, 2N)$ we have $U_N(\Phi\Phi') = U_N(\Phi)U_N(\Phi')$ (see (Mezzadri, 2002; Gurevich and Hadani, 2003) for improvements on the congruence condition). Consequently, if $\Phi\Phi' = \Phi'\Phi \bmod 2N$ then their propagators commute. This is the basic principle that we use to form the Hecke operators.

REMARK. The congruence $\Phi\Phi' = \Phi'\Phi \bmod 2N$ is much less restrictive than the equation $\Phi\Phi' = \Phi'\Phi$. The latter has as its solutions in $\mathrm{SL}(2, \mathbb{Z})$ essentially only ± powers of Φ (at least for Φ "primitive").

5.2. EQUIDISTRIBUTION OF HECKE EIGENFUNCTIONS

Since the Hecke operators commute with $U_N(\Phi)$, they act on its eigenspaces, and since they commute with each other there is a basis of $\mathcal{H}_N$ consisting of joint eigenfunctions of $U_N(\Phi)$ and the Hecke operators, whose elements we call Hecke eigenfunctions.

THEOREM 5.1 ((Kurlberg and Rudnick, 2000)). *Let $\Phi \in \mathrm{SL}(2,\mathbb{Z})$ be hyperbolic, $\Phi = I \bmod 4$, and $f \in C^\infty(\mathbb{T})$ a smooth observable. Then for all normalized Hecke eigenfunctions $\psi \in \mathcal{H}_N$ of $U_N(\Phi)$, the expectation values $\langle \mathrm{Op}_N(f)\psi, \psi\rangle$ converge to the phase-space average of f as $N \to \infty$. Moreover, for all $\epsilon > 0$ we have*

$$\langle \mathrm{Op}_N(f)\psi, \psi\rangle = \int_{\mathbb{T}} f(x)\,dx + O_{f,\epsilon}(N^{-1/4+\epsilon}), \quad \text{as } N \to \infty.$$

The exponent of $\frac{1}{4}$ in this theorem is not optimal. What is in fact shown is that if ψ_i, $i = 1, \ldots, N$ is an orthonormal basis of $\mathcal{H}_N$ consisting of Hecke eigenfunctions then

$$\sum_{i=1}^{N} \left| \langle \mathrm{Op}_N(f)\psi_i, \psi_i\rangle - \int_{\mathbb{T}} f(x)\,dx \right|^4 \ll N^{-1+\epsilon} \tag{17}$$

(compare Theorem 3.2). We deduce Theorem 5.1 from (17) by taking an orthonormal basis with $\psi_1 = \psi$ and omitting all but one term on the LHS. If all terms on the LHS of (17) are of roughly the same size then we would expect this to give the exponent $\frac{1}{2}$.

For prime values of N, this is in fact known; for primes where Φ is diagonalizable modulo N (this is 50% of the primes), this is shown in (Degli Esposti et al., 1995), while for the remaining 50% of primes this was shown by Gurevich and Hadani (Gurevich and Hadani, 2006). A key ingredient here is the Weil bound on complete exponential sums.

5.3. HIGHER DIMENSIONS

There is an extension of the above theory for linear symplectic maps of higher dimensional tori, including the existence of Hecke operators. Kelmer (Kelmer, 2005) discovered that the analogue of arithmetic quantum unique ergodicity fails in certain cases, and has shown that this happens if and only if there are isotropic rational subspaces, invariant under the linear map. In the case that there are invariant isotropic rational subspaces, Kelmer constructed "super-scars": For each prime value of N, there are Hecke eigenfunctions $\psi_N \in \mathcal{H}_N$ that localize on a co-isotropic invariant subtorus.

References

Bourgain, J. (2006) A remark on quantum ergodicity for cat maps, preprint.

Bouzouina, A. and De Bièvre, S. (1996) Equipartition of the eigenfunctions of quantized ergodic maps on the torus, *Comm. Math. Phys.* **178**, 83–105.

Colin de Verdière, Y. (1985) Ergodicité et fonctions propres du laplacien, *Comm. Math. Phys.* **102**, 497–502.

De Bièvre, S. (2006) An introduction to quantum equidistribution, in this book.

Degli Esposti, M., Graffi, S., and Isola, S. (1995) Classical limit of the quantized hyperbolic toral automorphisms, *Comm. Math Phys.* **167**, 471–507.

Faure, F., Nonnenmacher, S., and De Bièvre, S. (2003) Scarred eigenstates for quantum cat maps of minimal periods, *Comm. Math. Phys.* **239**, 449–492.

Friedlander, J. (2006) Uniform distribution, exponential sums and crypography, in this book.

Gurevich, S. and Hadani, R. (2003) The two dimensional Hannay–Berry model, arXiv:math-ph/0312039.

Gurevich, S. and Hadani, R. (2006) Proof of the Kurlberg–Rudnick Rate Conjecture, *C. R. Math. Acad. Sci. Paris* **342**, 69–72.

Hannay, J. H. and Berry, M. V. (1980) Quantization of linear maps on a torus - Fresnel diffraction by a periodic grating, *Physica D* **1**, 267–291.

Kelmer, D. (2005) Arithmetic quantum unique ergodicity for symplectic linear maps of the multidimensional torus, arXiv:math-ph/0510079.

Kurlberg, P. (2003) On the order of unimodular matrices modulo integers, *Acta Arith.* **110**, 141–151.

Kurlberg, P. and Rudnick, Z. (2000) Hecke theory and equidistribution for the quantization of linear maps of the torus, *Duke Math. J.* **103**, 47–78.

Kurlberg, P. and Rudnick, Z. (2001) On quantum ergodicity for linear maps of the torus, *Commun. Math. Phys.* **222**, 201–227.

Lindenstrauss, E. (2006) Three examples how to use measure classification in number theory, in this book.

Mezzadri, F. (2002) On the multiplicativity of quantum cat maps, *Nonlinearity* **15**, 905–922.

Rudnick, Z. and Sarnak, P. (1994) The behaviour of eigenstates of arithmetic hyperbolic manifolds, *Comm. Math Phys.* **161**, 195–213.

Schnirelman, A. (1974) Ergodic properties of eigenfunctions, *Usp. Math. Nauk* **29**, 181–182.

Zelditch, S. (1987) Uniform distribution of eigenfunctions on compact hyperbolic surfaces, *Duke Math. J.* **55**, 919–941.

Zelditch, S. (1996) Quantum ergodicity of C^*-dynamical systems, *Comm. Math.Phys.* **177**, 507–528.

Zelditch, S. (1997) Index and dynamics of quantized contact transformations, *Ann. Inst. Fourier (Grenoble)* **47**, 305–363.

INDEX